

# NETWORK SYNTHESIS BY IMPULSE RESPONSE FOR SPECIFIED INPUT AND OUTPUT IN THE TIME DOMAIN

FREDDY BA HLI

LOAN COPY mly

TECHNICAL REPORT NO. 261

JULY 31, 1953

RESEARCH LABORATORY OF ELECTRONICS  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
CAMBRIDGE, MASSACHUSETTS

The Research Laboratory of Electronics is an interdepartmental laboratory of the Department of Electrical Engineering and the Department of Physics.

The research reported in this document was made possible in part by support extended the Massachusetts Institute of Technology, Research Laboratory of Electronics, jointly by the Army Signal Corps, the Navy Department (Office of Naval Research), and the Air Force (Office of Scientific Research, Air Research and Development Command), under Signal Corps Contract DA36-039 sc-42607, Project 132B; Department of the Army Project 3-99-12-022; and in part by a fellowship granted to the author under the technical aid program of GUB-ECA (transferred later to GUB-TCA).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
RESEARCH LABORATORY OF ELECTRONICS

Technical Report No. 261

July 31, 1953

NETWORK SYNTHESIS BY IMPULSE RESPONSE FOR SPECIFIED  
INPUT AND OUTPUT IN THE TIME DOMAIN

Freddy Ba Hli

This report is based on a thesis submitted to the Department of Electrical Engineering, M.I.T., in partial fulfillment of the requirements for the degree of Doctor of Science, 1953.

Abstract

A solution to the general problem of synthesis in the time domain is presented: Given an arbitrary time function as the specified input to an electrical network and another such time function as the desired output, we offer a design procedure for a network for obtaining the required relationship. These arbitrary time functions may be prescribed either graphically, or as analytic functions, or merely as sequences of values at stated intervals of time.

The specified input-output pair is reduced to a single function which is known as the impulse response of the network, by a simple, straightforward, synthetic division procedure. The physical realizability of the network is tested by two, easy-to-apply, criteria in the time domain that are the analogs of the Hurwitz test used on functions of  $s$ .

Then three original methods are presented for calculating the network system function directly from the impulse response in the time domain. All three methods offer the advantages of being simple and computationally rapid. Moreover, it is possible to get rational function approximations to transcendental network system functions by an algebraic method which yields great economy of network elements with good tolerances in the time domain.



# NETWORK SYNTHESIS BY IMPULSE RESPONSE FOR SPECIFIED INPUT AND OUTPUT IN THE TIME DOMAIN

## I. INTRODUCTION

### 1.1 TERMINOLOGY

Up to a few years ago, network synthesis concerned itself almost exclusively with the design of electrical networks of the linear, passive variety for obtaining some required behavior as a function of frequency. This was an outcome of the predominant development of network theory by the telephone, telegraph, and radio engineers for designing newer and better filters to pass or reject certain bands of frequencies. Carrier telephony, telegraphy, and the increasing use of the radio spectrum all contributed to the diversity and complexity of this problem.

With the advent of television, radar, pulse modulation, servomechanisms, high-speed electronic computers, and a host of other new devices in communications engineering and in various other fields too, the emphasis has shifted recently to what is called the transient problem. This does not mean that the problem is transitory; on the contrary, it will probably stay with us for quite some time because it is extremely difficult! What is meant is that we wish to consider the problem from the aspect of time functions or transient phenomena.

The term "synthesis in the time domain" will be employed to distinguish this field from the usual notion of network synthesis for prescribed frequency behavior. Conventional network synthesis will not, therefore, be alluded to, but it is implicitly understood that the ultimate production of an electrical network will have to make use of the well-known theorems and methods of that branch of electrical communications engineering.

For the benefit of those to whom this new field of "synthesis in the time domain" is unfamiliar, we add a few more remarks before giving a statement of the general problem. Because we are dealing with passive networks, both the input and the output of the network are generally specified. In most cases we cannot do much about the input; presumably it comes to us as a television or radar signal or from some other device and we are to try to make a network that will produce the desired output for that input. As a result of the lack of any general synthesis procedures using time functions, the method of solution has been to utilize some synthesis procedure from the large stock of frequency-domain methods and to hope for the best in the time domain, namely that it will be a reasonable approximation to what is wanted. For example, consider the simple case of getting an amplifier that will have good transient response, that is, will amplify sharp pulses without too much distortion. One procedure would be to determine the equivalent bandwidth in frequency corresponding to the duration of the pulses, and then to go ahead and design an amplifier whose frequency response is flat within that bandwidth.

With the popularity of "hi-fi," even the layman can now talk in terms of frequency

response characteristics and harmonic distortion. So, although the phenomena we are interested in are readily expressed in the time domain, we usually talk of them in terms of frequency functions. Now the basic concept of a frequency as so many cycles per second implies a steady-state condition. Provided that the things that are happening have settled down to the steady state, this approach is neat and easy because, since the time of Steinmetz and his introduction of the methods of complex algebra to sinusoidal steady-state analysis, this whole field has been well catalogued and the solutions and basic theorems are well known.

The sinusoid, which is so familiar to steady-state, alternating-current workers, lets itself be handled without much trouble with complex algebra or with rotating vectors but is as hard to handle as a snake, which it resembles, in the time domain. This factor, among others, has contributed to the belief that synthesis in the time domain is a difficult proposition.

## 1.2 STATEMENT OF THE PROBLEM

Briefly, we may state that the problem to be tackled in this research is to synthesize an electrical network that will produce a specified function of time  $f_o(t)$  as its output when excited by a prescribed input function of time  $f_i(t)$ .

The use of the words "function of time" does not necessarily mean that they are given to us as analytic functions. The word "function" is used in its broad mathematical sense as being some relation between the dependent quantity  $f(t)$  and the independent variable  $t$ . So the input or output may be specified as a curve plotted with  $f(t)$  as ordinate and  $t$  as abscissa. Or they may be in tabular form, giving numerical values of  $f(t)$  at stated intervals of time.

It is not proposed in this introduction to enter deeply into the solution of the general problem of time-domain synthesis; but it may readily be appreciated that since we are given a pair of functions, and the network to do this presumably has one function which describes it uniquely, the first step would be to obtain this describing function. From the Laplace transform theory we do know a function  $h(t)$  that relates the input and output of a network. It is called the impulse response of the network, and its Laplace transform is appropriately called the network system function because from this function we can synthesize the network. However, this is not at all the engineering solution, nor even a complete mathematical solution, because of the distressing fact that  $f_i(t)$  and  $f_o(t)$  are related by  $h(t)$  through a finite integral known as the convolution integral in the English texts. This means that in order to get  $h(t)$  from

$$f_o(t) = \int_0^t f_i(\tau) h(t-\tau) d\tau \quad (1)$$

we have to solve an integral equation, and this, in general, is something that an engineer likes to avoid. It may turn out that even the mathematician is not able to get a solution

to the integral equation given above. Furthermore, if  $f_i(t)$  and  $f_o(t)$  are not given as analytic functions, Eq. 1 is not available.

We searched for a method that would enable us to calculate  $h(t)$  under all conditions. Having obtained a simple, straightforward procedure that produces an accurate impulse response to within the tolerances set by the needs of the particular case we are considering, the main problem was to get the system function  $H(s)$  corresponding to  $h(t)$  as a rational function. We know from experience that many input-output pairs that we can solve to get  $H(s)$  by Laplace transforms come out with  $H(s)$  that are transcendental functions; that is, they either have an infinite number of poles in the  $s$  plane or have an essential singularity at infinity. Transcendental functions, therefore, lead to network designs with an infinite number of lumped-parameter elements or with distributed parameters. Because rational functions contain only a finite number of poles, they lead to finite, lumped-parameter networks, which is what we would like to have. This is an important problem on which much research work has been expended, as the references will show. Furthermore, for the purpose of our investigation we need to get the  $H(s)$  directly from a curve or a sequence of values of  $h(t)$  at stated intervals of time because that may be the form in which the impulse response is given to us. In addition, we would prefer to get this rational function approximation (for those cases where  $h(t)$  corresponds to a transcendental Laplace transform) with as few poles as possible for any specified value of tolerance because this means an economy of network elements. These are the various problems to be solved and the desired conditions on the solutions.

## II. HISTORICAL SURVEY

### 2.1 REVIEW OF RESEARCH ON THE PROBLEM OF TIME-DOMAIN SYNTHESIS AND THE REASONS FOR THE EMERGENCE OF THIS PROBLEM

It was mentioned in section 1.1 that network synthesis in the past has generally referred only to the frequency domain. During World War II, the tremendous development of radar and computing devices, and servomechanisms brought out the importance of considering the effects of transients in all kinds of electrical devices, and a need arose for designing networks that would have prescribed transient behavior. With Coulthard, we may say that communication is just the science of controlled transients. For if the signal has settled down to the steady state, it does not convey any more information except to indicate that it is still there! After the war, when television started really to expand, electrical networks with more stringent transient performance were demanded. The reason for this is that although the ear is very tolerant to phase distortion, the eye is not. Hence any network that transmits information with care for only the amplitudes of the various frequency components, and lets the phase come as it may, is not at all satisfactory for television programs, although it may be quite adequate for radio programs. This talk in terms of phase and amplitude of different frequency components will be at once recognized as a legacy from the filter theory days. Actually, the phenomenon is in the time domain; that is, the input is a function of time, the output is a function of time, and the network itself is an operator that acts on the input to produce the output. Because of the difficulty of working mathematically with these operations in the time domain, we find that the first reports on these problems preferred to work with the frequency-domain concepts. This is quite understandable because there was nothing well formulated in terms of time-domain operators, whereas there was a vast wealth of procedures in terms of functions of  $s$  (refs. 1-9).

The only available texts on time-domain problems were on transient analysis (refs. 10-14); that is, given a network, they show how to find the transient behavior; but given a transient behavior, they do not show how to find the network.

A listing of the theses and reports indicates the influence of the modern electrical devices exerting themselves by asking for new methods of synthesis of pulse generators, television amplifiers, and feedback networks. For example, Fawwaz (16) was looking for pulse-forming networks (21), (22); Marsten's problem (17) was a video amplifier; Bond (18) was working on the synthesis of servomechanisms (19). Marsten's problem, though not solvable with the techniques that he had available in 1946, may become tractable by the use of new methods of Weinberg (23) for synthesis of transfer functions of  $s$  into RLC and RC networks, developed in 1951.

Tuttle (24), in 1948, discussed many methods of approximating the time function that is the desired impulse response by means of a sum of damped exponentials. For least-mean-square approximation the orthogonal functions, such as the Laguerre polynomials, predistorted exponentials, and Jacobi polynomials may be tried; for Chebyshev behavior



those polynomials called pulse functions are useful. The objections to the orthogonal function approximation method from a practical standpoint are said to be impressive, and the method is not recommended for network synthesis, since least-mean-square error is not necessarily desirable although mathematically convenient. Taylor approximations and Prony's method by exponentials, Tuttle says, require considerable computation. The Lagrangian interpolation coefficients give speedy approximations with a restricted type of exponentials, but the method is of the cut-and-try category and therefore not elegant. Tuttle had hopes for the Prony technique by employing some extension of it and Carr (25) carried it out. It is actually an orthogonal approximation of a function in terms of its derivatives or integrals, and a systematic procedure is obtained for both pole location and coefficient value. It suffers from a lack of useful criteria for determining beforehand whether or not the ensuing network will be realizable.

White (26) was the next to take up the subject. He came out with his method of transitional transients that are generated by  $W(s, t)$  in the transform

$$f(t) = \frac{1}{2\pi j} \int F(s) \exp [W(s, t)] ds \quad (2)$$

that occurs in the saddlepoint method of integration.

The latest contributor is Kautz (27), who also started out from the saddlepoint method and developed his constituent transients approach. This employs, as time-function approximants to the desired impulse response, certain eigenfunctions (such as the set of Bessel functions) that are capable of accurate and controlled approximation by frequency-domain expansions that lead to realizable system functions, such as the Padé functions. He also investigated the Chebyshev, Legendre, Hermite, and Laguerre polynomials and the group of lambda functions and the hypergeometric series as approximating functions.

The most significant contributions to the general theory of synthesis in the time domain are contained in the basic existence theorems set forth in "Rational Function Approximations for Network Functions" by Dr. M. V. Cerrillo and Professor E. A. Guillemin (28).

## 2.2 A SET OF BASIC EXISTENCE THEOREMS IN THE THEORY OF SYNTHESIS IN THE TIME DOMAIN FROM CERRILLO AND GUILLEMIN

These theorems are obtained by a set of new integral representations for the direct and inverse Laplace transformations, which are

$$F(s) = \frac{s(s - \gamma_0)}{\pi} \int_{\Gamma_0} \frac{U(\gamma_0, \lambda)}{(s - \gamma_0)^2 + \lambda^2} d\lambda \quad (3)$$

where  $s = \sigma + j\omega$ ,  $S = \gamma + j\lambda$ , and

$$f(t) = \frac{2 \exp(\gamma_0 t)}{\pi} \int_{\Gamma_0} U(\gamma_0, \lambda) \cos \lambda t \, d\lambda \quad (4)$$

where  $F(S_0) = U(\gamma_0, \lambda) + jV(\gamma_0, \lambda)$  for  $\gamma_0 > c_0$  and the particular contour  $\Gamma_0$ .

For the construction of the rational expansions of transfer functions, these integrals need to be generalized so that even if they do not exist in the Riemann sense a representation can be obtained by the Stieltjes integrals. To this end, the distribution functions are introduced, defined by

$$\phi(\gamma_0, \lambda) = \int_0^\lambda U(\gamma_0, \mu) \, d\mu \quad (5)$$

and

$$\tau(\gamma_0, t) = \int_0^t f(\mu) \exp(-\gamma_0 \mu) \, d\mu \quad (6)$$

These functions exist if  $f(t)$  is Laplace-transformable. Using them, the integrals, Eqs. 3 and 4, become

$$F(s) = \frac{2}{\pi} (s - \gamma_0) \int_{\Gamma_0} \frac{d\phi(\gamma_0, \lambda)}{(s - \gamma_0)^2 + \lambda^2} \quad (7)$$

$$f(t) = \frac{2 \exp(\gamma_0 t)}{\pi} \int_{\Gamma_0} \cos \lambda t \, d\phi(\gamma_0, \lambda) \quad (8)$$

Then, if we define a transfer function as the difference of two positive-real functions, the fundamental theorem of the existence of transfer functions is obtained as:

Theorem I

"The necessary and sufficient condition for a function  $F(s)$  to be a transfer function is that it can be represented by the Stieltjes integral

$$F(s) = \frac{2(s - \gamma_0)}{\pi} \int_{\Gamma_0} \frac{d\phi(\gamma_0, \lambda)}{(s - \gamma_0)^2 + \lambda^2} \quad (9)$$

or

$$\frac{2s}{\pi} \int_0^\infty \frac{d\phi(0, \lambda)}{s^2 + \lambda^2} \quad (10)$$

for every contour  $\Gamma_0$  that shall be made to coincide with the upper part of the imaginary axis as indicated by the last integral. Here  $\phi(\gamma_0, \lambda)$  and  $\phi(0, \lambda)$  are functions of bounded

variation, that is

$$\int_0^{\infty} |U(\gamma_0, \lambda)| d\mu < \infty \quad \text{and} \quad \int_0^{\infty} |U(0, \lambda)| d\mu < \infty."$$

Furthermore, two corollaries are derived from Theorem I.

#### Theorem II

"Let  $f(t)$  be a real, single-valued, bounded-almost-everywhere function, which is zero for  $t < 0$ ;  $f(t)$  may possess a denumerable set of isolated points of simple discontinuity. Then its Laplace transform is necessarily a transfer function."

#### Theorem III

"Let  $F(s)$  be a general transfer function as defined above. Then its inverse Laplace transform is necessarily a function  $f(t)$  with the properties listed in Theorem II (except for a set of zero measure)."

These important theorems define the open field of network synthesis possibilities.

Dr. Cerrillo, F. Bolinder, and Ba Hli worked on a method for calculating the rational function approximation for some types of time functions that may be desired as impulse responses. It depended on Eq. 3 and used the real part of  $F(s)$  evaluated along the imaginary axis,  $s = j\omega$ , from  $-\infty$  to  $+\infty$ . This real part was then approximated by a number of zeros and poles to form a rational function, the pole locations being determined by use of van der Monde determinants and symmetric functions.

Dr. Cerrillo and Ba Hli also reported another method that employed the same principle of approximating the real part of  $F(s)$  corresponding to a time function, along the imaginary axis, but performed this approximation by a different technique. (Quarterly Progress Report, Research Laboratory of Electronics, M.I.T., Jan. 15, 1953).

The germ of the idea for what is called the iterative substitution method for obtaining the impulse response from specified input and output came from an article by M. F. M. Osborne of the Naval Research Laboratories, Washington, D. C. (J. Appl. Phys. 14, 180, 1943). It concerns a method of transient analysis of linear systems based on Duhamel's theorem in the Heaviside unit-step function theory:

$$i(t) = e(0) A(t) + \int_0^t A(t-\lambda) e'(\lambda) d\lambda \quad (11)$$

where  $i(t)$  is the response to an arbitrary driving force  $e(t)$ , and  $A(t)$  is the unit-step response. To those who know the relation between Heaviside operational calculus and the Laplace transformation method, Eq. 11 will appear as the analog of the convolution integral, since the unit step is 1 for Heaviside, while the unit impulse has the Laplace transform 1. In the next chapter we shall see this method developed.

### III. IMPULSE RESPONSE FROM INPUT AND OUTPUT

#### 3.1 METHOD OF OBTAINING THE IMPULSE RESPONSE OF THE NETWORK

From the convolution integral (also called the Faltung after the German term) that relates the output  $f_o(t)$  to the input  $f_i(t)$

$$f_o(t) = \int_0^t f_i(\tau) h(t-\tau) d\tau \quad (12)$$

it is seen that the function  $h(t)$  is all that is needed to be able to write  $f_o(t)$  when  $f_i(t)$  is given. In the language of network theory, this means that since a particular network produces  $f_o(t)$  when excited by  $f_i(t)$ ,  $h(t)$  specifies this network completely in the time domain. Hence the complex-frequency transform  $H(s)$  in turn specifies the network completely in the complex frequency domain, and, therefore, it is known as the network system function.

In terms of functions of  $s$ , we then have

$$F_o(s) = F_i(s) \cdot H(s) \quad (13)$$

If  $F_i(s) = 1$ , which means that the excitation is an impulse function in time, then the output is just  $h(t)$ . Therefore,  $h(t)$  is the impulse response of the network. Thus, from the preceding paragraph, it is deduced that if we can find  $h(t)$  from given  $f_i(t)$  and  $f_o(t)$ , then we have gone a big step forward towards the synthesis of the network, since the network system function  $H(s)$  can then be determined.

Equation 13 is very simple in contrast to Eq. 12. Moreover, it is very familiar to networks engineers. For them, Eqs. 13 and 12 are the convenient formulation of the theorem on complex multiplication in the Laplace transform theory (12), which states:

"Convolution in the real (time) domain goes over into multiplication in the complex (frequency) domain."

The theorem provides a good example of how the Laplace transformation converts a complicated operation (convolution) in the real domain of  $t$ , into a simpler operation (multiplication) in the complex frequency domain of  $s$ .

The question now presents itself: Is it possible to retain this simplicity of operation while still using the values of the functions in the time domain? We ask this because, in synthesis for specified time behavior, we would like to operate as much as possible and perform whatever approximations we may need to get an engineering solution to the problem in the time domain, since our experience with approximations in  $s$  teaches us that these sometimes yield very poor results in  $t$ .

If we look at Eq. 12, we readily appreciate that to work with  $f_i(t)$  and  $f_o(t)$  involves solving an integral equation that does not appear to be a simple operation at all. This leads to the feeling that time-domain synthesis should be left alone, and that further work should be in getting better approximations in the complex frequency domain.

However, the answer to the above question is really in the affirmative. The secret is to replace the integral by a summation. This may, of course, at once stir up prejudice against the method, because we feel that it may be a nonconverging approximation. But this need not always be so and it is shown, both theoretically and in practice, that consistent and accurate results are obtainable. In this section we shall describe the method and in the later sections the theoretical and practical justification. In reading through this portion, one should bear in mind that this is mainly a description and that the full discussion is postponed, because we feel that a clearer understanding would result if a picture of the general procedure were available.

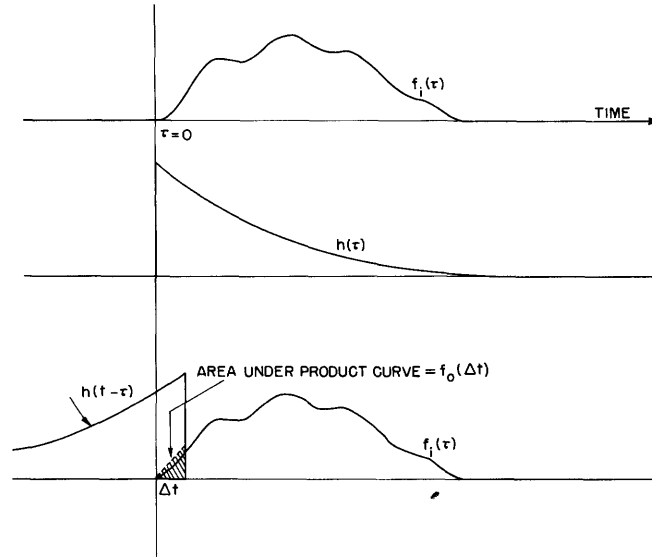


Fig. 1

The basic idea is simple and merely makes use of Leibnitz and Newton's fundamental conception of an integral as the limit of a summation of areas. If we examine the details of the convolution integral, Eq. 12, in this light, we see that the first value of  $f_o(t)$  at say  $t = \Delta t$ , where  $\Delta t \rightarrow 0$ , is given by (a) multiplying  $f_i(t)$  by  $h(t)$  folded about the  $t = 0$  axis, and moved to the right by  $\Delta t$  and (b) taking the area under the product curve (12).

Written out this would be

$$f_{o_1} = f_{i_1} \Delta \tau h_1 \quad (14)$$

where  $f_{o_1}$  = value of  $f_o(t)$  at  $t = \Delta t$ ,

$f_{i_1}$  = value of  $f_i(t)$  at  $t = \Delta t$ ,

$h_1$  = value of  $h(t)$  at  $t = \Delta t$ ,

carrying through an analogous procedure for  $t = 2\Delta t$ ,  $3\Delta t$ , and so on,

$$\begin{aligned} f_{o2} &= f_{i1} \Delta\tau h_2 + f_{i2} \Delta\tau h_1 \\ f_{o3} &= f_{i1} \Delta\tau h_3 + f_{i2} \Delta\tau h_2 + f_{i3} \Delta\tau h_1 \end{aligned} \quad (15)$$

Now we may solve for  $h_1$ ,  $h_2$ , and the like, as

$$\begin{aligned} h_1 &= \frac{f_{o1}}{f_{i1} \Delta\tau} \\ h_2 &= \frac{f_{o2} - f_{i2} \Delta\tau h_1}{f_{i1} \Delta\tau} \\ &\vdots \end{aligned} \quad (16)$$

and in general

$$h_n = \frac{f_{on} - \sum_{p=2}^n f_{ip} \Delta\tau h_{n+1-p}}{f_{i1} \Delta\tau}$$

Because of the nature of the process, we may call this an iterative substitution method. The sequence of values of  $h$  will be a sequence in time and henceforth will be denoted by

$$\{h_1, h_2, h_3, \dots, h_n, \dots\}$$

or simply by  $\{h\}$  for brevity. It is a representation in sequence form for  $h(t)$ .

The importance of  $h(t)$  has been brought out earlier as specifying the network in time, and through its Laplace transform  $H(s)$ , as specifying completely the network system function.

We see that the above procedure will enable us to calculate by simple straightforward numerical methods the impulse response of the network required to produce the prescribed  $f_o(t)$  from  $f_i(t)$ . It will be shown in section 3.3 that we may obtain  $h(t)$  to within as close a tolerance as we wish by a suitable choice of  $\Delta t$  for subdividing the time axis of the  $f(t)$ .

The next development is to further reduce the computations by recognizing that the process of getting  $h_1, h_2, h_3, \dots$ , is just the same as that which we would go through in calculating the quotient of two polynomials where the  $f_{in}$  and the  $f_{on}$  are the coefficients of like powers in the two polynomials. There is a very definite theoretical reason for this that will be given in section 3.2, but for the time being let us write  $f_o(t)$  and  $f_i(t)$

as time sequences. That is, divide time into equal intervals  $\Delta t$ , and write the values of  $f_o(t)$  at  $t = \Delta t, 2\Delta t, 3\Delta t, \dots$

$$\{f_{o_1}, f_{o_2}, f_{o_3}, \dots\}$$

Similarly for  $f_i(t)$

$$\{f_{i_1}, f_{i_2}, f_{i_3}, \dots\}$$

A point of detail now arises. How big or how small can we choose  $\Delta t$ ? Obviously if we took a  $\Delta t$  so wide that either  $f_i(t)$  or  $f_o(t)$  went through a whole cycle of oscillation within that period, then the time sequences would not be true indications of the behavior of  $f(t)$ . On the upper limit we may say that  $\Delta t$  must be small enough so that the time sequences will indicate the behavior of  $f(t)$  adequately. For example, if  $f(t)$  has an oscillation within a time period of one second, then  $\Delta t$  should be small enough so that at least two or three values of  $f(t)$  should be picked up on the ascending or descending part of the curve before the first maximum or minimum.

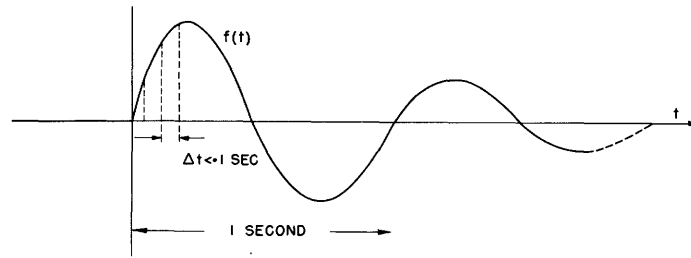


Fig. 2

On the other hand, as we let  $\Delta t \rightarrow 0$ , the number of terms in the sequence will mount up, which will increase the computations. Also the successive values of  $f(t)$  as  $\Delta t \rightarrow 0$  become less and less different, and so for most cases, a good idea of what  $h(t)$  is would be obtained by choosing  $\Delta t$  such that both  $f_i(t)$  and  $f_o(t)$  change on the average by about 10 percent of their maximum values within successive intervals. An important point is that by picking first a rather wide interval for  $\Delta t$  (which is within the upper limit prescribed above) and plotting  $\{h\}$ , which we shall denote by  $\{h\}_1$  corresponding to  $\Delta t = \Delta t_1$ ; then if we use a  $\Delta t = (\Delta t_1/2), (\Delta t_1/4)$  and so on, we get successive  $\{h\}_2, \{h\}_3$  which converge rapidly towards a limiting  $\{h\}$ .

Having chosen a convenient value of  $\Delta t$ , to obtain the two sequences  $\{f_o\}$  and  $\{f_i\}$ , we then perform synthetic division to get  $\{h\}_A$  as

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}} \quad (17)$$

where  $\{h\}_A$  denotes the sequence of areas under the impulse response curve. In other words, the sequence  $\{h\}_A$  would consist of terms  $h_1\Delta t$ ,  $h_2\Delta t$ ,  $\dots$ .

If we carry out the synthetic division we get

$$\begin{array}{r}
 \frac{f_{o1}}{f_{i1}}, \left( f_{o2} - \frac{f_{i2}f_{o1}}{f_{i1}} \right) / f_{i1}, \dots \\
 \hline
 f_{i1}, f_{i2}, f_{i3}, \dots \quad \cancel{f_{o1}}, f_{o2}, f_{o3}, \dots \\
 \quad \quad \quad \cancel{f_{o1}}, \frac{f_{i2}f_{o1}}{f_{i1}}, \frac{f_{i3}f_{o1}}{f_{i1}}, \\
 \hline
 \left( f_{o2} - \frac{f_{i2}f_{o1}}{f_{i1}} \right), \left( f_{o3} - \frac{f_{i3}f_{o1}}{f_{i1}} \right), \dots \\
 \left( f_{o2} - \frac{f_{i2}f_{o1}}{f_{i1}} \right), \left( f_{o2} - \frac{f_{i2}f_{o1}}{f_{i1}} \right) \frac{f_{i2}}{f_{i1}}, \dots
 \end{array}$$

When we compare the quotient sequence

$$\frac{f_{o1}}{f_{i1}}, \frac{\left( f_{o2} - \frac{f_{i2}f_{o1}}{f_{i1}} \right)}{f_{i1}}$$

with the set of formulae, Eq. 16, we see that these terms in the quotient sequence are identical to  $h_1\Delta t$ ,  $h_2\Delta t$ ,  $\dots$ , thereby giving us a more compact procedure for getting the impulse response.

### 3.2 THEORY OF THE METHOD; THE LAPLACE-STIELTJES INTEGRAL

The time has now come to lend a little theoretical support to the procedure. The theoretical structure can be obtained readily from the calculus of finite differences; more particularly, the one called symbolical calculus. This differs from the infinitesimal calculus, with which we usually work, in that it is applicable to both continuous variables and discontinuous variables, whereas the infinitesimal calculus is restricted to only continuous variables.

In symbolical calculus the notation  $E f(t)$  is used by the English writers (30), Shepard, Milne-Thomson (31), Whittaker and Robinson, for what is called the operation of displacement. The actual operation was introduced by Boole and was denoted as  $D f(t)$  by him. De la Vallee Poussin also introduced it and denoted it as  $\nabla f(t)$ . Since  $D$  and  $\nabla$  now have gained wide acceptance for other operators, the symbol  $E$  is employed instead.



The displacement operator  $E$  has the property that:

$$\begin{aligned} \text{a. } & Ef(t) = f(t + \Delta t) \\ & E^2 f(t) = Ef(t + \Delta t) = f(t + 2\Delta t) \\ & E^3 f(t) = E^2 f(t + \Delta t) = Ef(t + 2\Delta t) = f(t + 3\Delta t) \end{aligned} \quad (18)$$

b. It is distributive:

$$E^n[f(t) + \phi(t) + \psi(t)] = E^n f(t) + E^n \phi(t) + E^n \psi(t) \quad (19)$$

c. It is commutative:

$$E^n E^m f(t) = E^m E^n f(t) = E^{m+n} f(t) \quad (20)$$

d. It behaves as an algebraic quantity with respect to addition, subtraction, division, and multiplication;

e. A polynomial in  $E$  also represents an operation. Several such polynomials may be added, subtracted, divided, and multiplied;

f. Negative powers of  $E$  are permitted and behave like the positive powers.

From these properties it is deduced that what has been called the output time sequence

$$\{f_{o_1}, f_{o_2}, f_{o_3}, f_{o_4}, \dots, f_{o_n}\}$$

is a polynomial in  $E$

$$f_{o_1} E + f_{o_2} E^2 + f_{o_3} E^3 + f_{o_4} E^4 + \dots + f_{o_n} E^n$$

operating on the unit-step function.

Similarly, the input time sequence

$$\{f_{i_1}, f_{i_2}, f_{i_3}, \dots, f_{i_m}\}$$

is another polynomial in  $E$

$$f_{i_1} E + f_{i_2} E^2 + f_{i_3} E^3 + \dots + f_{i_m} E^m$$

operating on the unit-step function.

It has been shown that polynomials in  $E$  may be treated as far as addition, subtraction, division, and multiplication are concerned, just like any algebraic polynomial. So this justifies the operation of synthetic division which we performed to get the areas under the impulse response curve because synthetic division is just a short-cut method for dividing one algebraic polynomial by another. Moreover, since the quotient of two polynomials is a third polynomial in the same symbol, we obtain as our quotient a polynomial in  $E$ , which is therefore equivalent to another time sequence, namely  $\{h\}_A$ .

We may show this result in another way. In order to do so we first need to consider the Laplace-Stieltjes integral

$$F(s) = \int_0^{\infty} \exp(-st) d\alpha(t) \quad (21)$$

This, as we will see, is a general integral which includes both the classical Laplace transform and the Dirichlet series. The connection between the Dirichlet series and the theory of the synthetic division method was brought out by Professor S. J. Mason. An analogous method was suggested by Professor W. K. Linvill during a thesis conference with Professor E. A. Guillemin.

Following Widder (32), we define a Stieltjes integral in the following manner:

Let  $\phi(x)$  and  $\psi(x)$  be real functions of the real variable  $x$  defined for  $a \leq x \leq b$ . Denote by  $\Delta$ , a subdivision of the interval  $(a, b)$  by the points  $x_0, x_1, \dots, x_n$  where

$$a = x_0 < x_1 < x_2 < \dots < x_n = b$$

By the norm  $\delta$  of  $\Delta$  we mean the largest of the numbers  $x_{k+1} - x_k$ ; ( $k = 0, 1, \dots, n-1$ ) Then if the limit

$$\lim_{\delta \rightarrow 0} \sum_{k=0}^{n-1} \phi(\mu_k) [\psi(x_{k+1}) - \psi(x_k)] \quad \text{where } x_k \leq \mu_k \leq x_{k+1}; (k = 0, 1, \dots, n-1)$$

exists independently of the manner of subdivision and of the choice of numbers  $\mu_k$ , then the limit is the Stieltjes integral of  $\phi(x)$  with respect to  $\psi(x)$  from  $a$  to  $b$ , namely

$$\int_a^b \phi(x) d\psi(x)$$

The Stieltjes integral reduces to the Riemann integral if  $\psi(x) = x$ .

The Stieltjes integral is more useful than the ordinary Riemann integral which is the one we generally employ in integral calculus, because when one of the functions is not continuous, the Stieltjes integral still exists, provided the functions are of bounded variation; this just implies the existence of the integrals

$$\int_0^{\infty} V[\phi(x)] dx < \infty \quad \text{and} \quad \int_0^{\infty} V[\psi(x)] dx < \infty$$

where  $V[\phi(x)]$  and  $V[\psi(x)]$  are the variations in  $\phi$  and  $\psi$ , which is true for all practical time functions that we would be interested in.

Hence, if we are dealing with step functions (which are discontinuous but of bounded variation), then the Stieltjes integral becomes a summation. In particular (33), if  $\psi(x)$

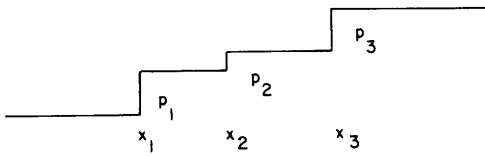


Fig. 3

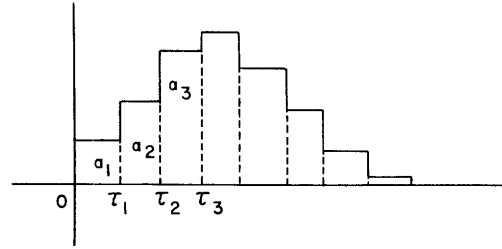


Fig. 4

is the step function with discontinuities  $p_1, p_2, p_3, \dots$  at the points  $x_1, x_2, x_3, \dots$ , the Stieltjes integral coincides with the sum

$$\sum p_k \phi(x_k)$$

which is a finite sum or an absolutely convergent infinite series according to whether or not the set of points of discontinuity is finite.

In the special case of the Laplace-Stieltjes integral, it becomes the Dirichlet series,

$$F(s) = \sum_{n=1}^{\infty} a_n \exp(-\tau) n^s \quad (22)$$

where the  $a_n$  are the areas under the  $f(t)$  curve during the intervals between  $\tau_n$  and  $\tau_{n-1}$ .

If  $f(t)$  is continuous, then we get the usual Laplace transform

$$F(s) = \int_0^{\infty} f(t) \exp(-st) dt \quad (23)$$

which exists as an integral in the Lebesgue sense.\*

Thus we have from Eqs. 21, 22, and 23 the complex frequency transforms for all the types of time functions that we shall be dealing with in network synthesis. Because of the generality of Eq. 21, we may take any time function for this purpose and represent its Laplace transform (by which we henceforth do not necessarily mean only the restricted form Eq. 23, but the more general one, Eq. 21) by a Dirichlet series merely by dividing the time base into intervals  $\Delta t$ , and taking the areas  $a_n$  under the curve during those intervals and forming the sum

$$F(s) = \sum a_n \exp(-\tau) n^s$$

---

\*For the distinction between Lebesgue integrals and Riemann integrals, we may refer to Titchmarsh (34); and for a few words on the notion of "measure," we may, for example, refer to Rogosinski(35).

where  $\tau_n = n\Delta t - (\Delta t/2)$  for this kind of subdivision. If  $\Delta t$  is sufficiently small, then  $\tau_n \approx n\Delta t$ .

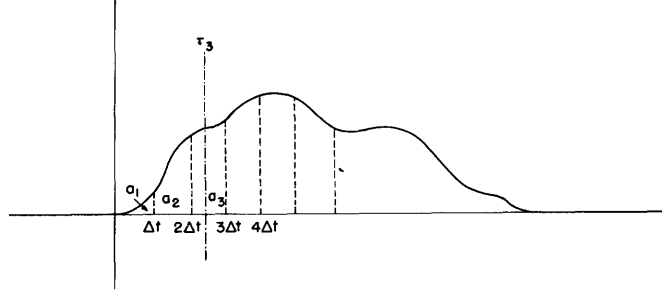


Fig. 5

Now we have the well-known result of network and Laplace theory, already given as Eq. 13, that

$$F_o(s) = F_i(s) \cdot H(s) \quad (24)$$

and therefore

$$H(s) = \frac{F_o(s)}{F_i(s)}$$

If we write  $F_o(s)$  and  $F_i(s)$  in the Dirichlet series form from the time functions  $f_o(t)$  and  $f_i(t)$ , we get what is essentially a ratio of two polynomials by treating  $\exp(-\Delta ts)$  as  $x$ . We can now see that  $H(s)$  would be given by synthetic division of  $F_o(s)$  by  $F_i(s)$ . For equal intervals  $\Delta t$ , the coefficients  $a_n$  in the Dirichlet series are just  $f_n \Delta t$ ; thus it is apparent that what we called our time sequences  $\{f_o\}$  and  $\{f_i\}$  are, except for the constant multiplier  $\Delta t$ , identical to the sequence of coefficients in the Dirichlet series in Eq. 24. Hence we obtained heuristically the true areas under the impulse response curve by our simple synthetic division in Eq. 17. The theoretical justification is now complete.

### 3.3 APPLICATION OF THE METHODS TO VARIOUS TYPES OF INPUTS AND OUTPUTS

We shall first work out an example with the iterative substitution method as given in Eq. 16, and then rework the same by the synthetic division method as in Eq. 17 for a comparison of the two methods of obtaining the impulse response.

(i) Let  $f_i(t)$  and  $f_o(t)$  be as in Fig. 6.

If we take the time intervals  $\Delta t$  as 0.2 (for a computation to illustrate the method

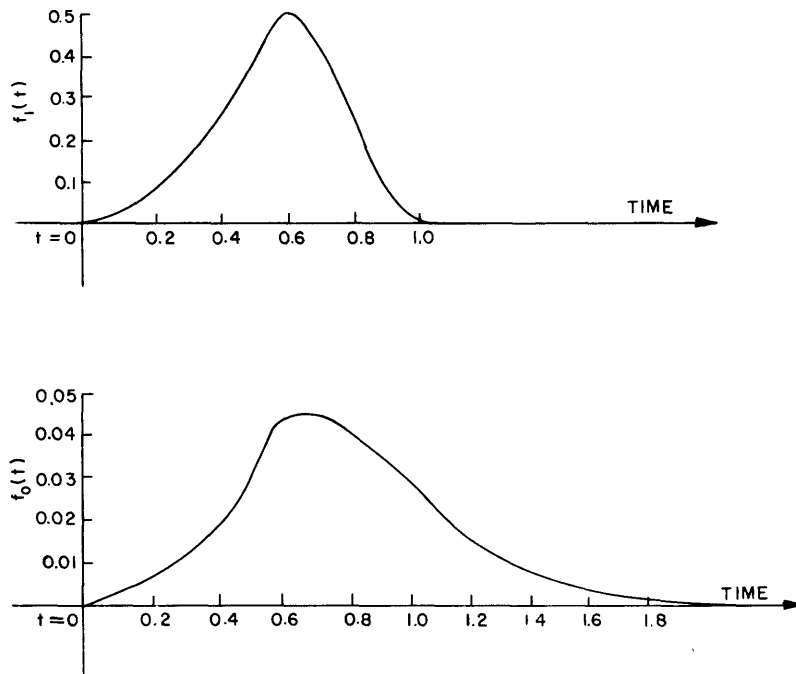


Fig. 6

with a minimum of arithmetic), we have

$$\{f_i\} = \{0.1, 0.25, 0.5, 0.25, 0, 0, \dots\}$$

$$\{f_o\} = \{0.0063, 0.0196, 0.0433, 0.0418, 0.0290, 0.0148, 0.0084, 0.0025, 0, 0, \dots\}$$

$$h_1 = \frac{f_{o1}}{f_{i1} \Delta t} = \frac{0.0063}{0.1 \times 0.2} = 0.315$$

$$h_2 = \frac{f_{o2} - f_{i2} \Delta t h_1}{f_{i1} \Delta t} = \frac{0.0196 - 0.25 \times 0.2 \times 0.315}{0.1 \times 0.2} = 0.19$$

$$h_3 = \frac{f_{o3} - f_{i3} \Delta t h_1 - f_{i2} \Delta t h_2}{f_{i1} \Delta t} = \frac{0.0433 - 0.5 \times 0.2 \times 0.315 - 0.25 \times 0.2 \times 0.19}{0.1 \times 0.2} = 0.11$$

$$h_4 = \frac{f_{o4} - f_{i4} \Delta t h_1 - f_{i3} \Delta t h_2 - f_{i2} \Delta t h_3}{f_{i1} \Delta t}$$

$$= \frac{0.0418 - 0.25 \times 0.2 \times 0.315 - 0.5 \times 0.2 \times 0.19 - 0.25 \times 0.2 \times 0.11}{0.1 \times 0.2} = 0.07$$

$$h_5 = \frac{f_{o5} - f_{i5} \Delta t h_1 - f_{i4} \Delta t h_2 - f_{i3} \Delta t h_3 - f_{i2} \Delta t h_4}{f_{i1} \Delta t}$$

$$= \frac{0.0290 - 0 - 0.25 \times 0.2 \times 0.19 - 0.5 \times 0.2 \times 0.11 - 0.25 \times 0.2 \times 0.07}{0.1 \times 0.2} = 0.05$$

$$h_6 = 0$$

If we use the synthetic division method we get

$$\begin{array}{r}
 \phantom{0.1, 0.25, 0.5, 0.25, 0, 0, \dots} \phantom{0.063, 0.038, 0.021, 0.014, 0.010, 0, 0, 0, 0, \dots} \\
 0.1, 0.25, 0.5, 0.25, 0, 0, \dots \left| \begin{array}{l}
 0.063, 0.038, 0.021, 0.014, 0.010, 0, 0, 0, 0, \dots \\
 0.0063, 0.0196, 0.0433, 0.0418, 0.0290, 0.0148, 0.0084, 0.0025, 0, \dots \\
 0.0063, 0.0158, 0.0316, 0.0158, \\
 \hline
 0.0038, 0.0117, 0.0260, 0.0290, \\
 0.0038, 0.0096, 0.0192, 0.0096, \\
 \hline
 0.0021, 0.0068, 0.0194, 0.0148, \\
 0.0021, 0.0054, 0.0150, 0.0054, \\
 \hline
 0.0014, 0.0044, 0.0094, 0.0084, \\
 0.0014, 0.0034, 0.0069, 0.0034, \\
 \hline
 0.0010, 0.0025, 0.0050, 0.0025, \\
 0.0010, 0.0025, 0.0050, 0.0025, \\
 \hline
 \vdots
 \end{array} \right.
 \end{array}$$

$$\{h\}_A = \{0.063, 0.038, 0.021, 0.014, 0.010, 0, 0, \dots\}$$

Since  $\Delta t = 0.2$ , the  $\{h\}_A$ , which is a sequence of areas, is converted at once to a sequence of ordinate values as

$$\{h\} = \{0.315, 0.19, 0.11, 0.07, 0.05, 0, 0, \dots\}$$

which is the same as the previous result by the less compact iterative substitution method. The impulse response is a simple exponential.

Any other example will yield identical results with both methods as may be seen from the law of formation of  $h_n$ ; consequently we will henceforth illustrate the synthetic division method only.

(ii) Next let us take the case of an ideal amplifier. This is particularly simple by the synthetic division procedure because

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}}$$

and for

$$\{f_o\} = k\{f_i\}$$

where  $k$  = amplification and is a scalar multiplier.

$$\{h\}_A = k\{1, 0, 0, 0, \dots\}$$

which is an approximation to an impulse of  $k$  units. Because of the finite size of  $\Delta t$ , the height of  $h(t)$  is not infinite but  $= (k/\Delta t)$ .

(iii) If we want a differentiator, let us specify, for simplicity, a unit ramp as our input. We could use more complicated functions without an increase in complexity of the impulse response, but the arithmetic would be tedious. Then let the output be

$$\{f_o\} = \{1, 1, 1, 1, 1, \dots\} \text{ the unit step.}$$

$$\{f_i\} = \{\Delta t, 2\Delta t, 3\Delta t, 4\Delta t, 5\Delta t, \dots\} \text{ unit ramp function.}$$

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}}$$

$$\begin{array}{r} \Delta t, 2\Delta t, 3\Delta t, 4\Delta t, 5\Delta t, 6\Delta t, \dots \left| \begin{array}{r} = \frac{1}{\Delta t}, -\frac{1}{\Delta t}, 0, 0, 0, 0, \dots \\ \hline 1, \quad 1, \quad 1, \quad 1, \quad 1, \quad 1, \dots \\ 1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6, \dots \\ \hline -1, -2, -3, -4, -5, \dots \\ -1, -2, -3, -4, -5, \dots \\ \hline 0, 0, 0, 0, 0, \dots \end{array} \right. \end{array}$$

This impulse response is well known theoretically as the unit doublet. Incidentally, this example is a good one to try the convolution of  $f_i(t)$  with  $h(t)$  because of the insight it affords us into the mechanism by which the time operator, known as the unit doublet, differentiates the time function it operates upon.

$$F_o(s) = F_i(s) \cdot H(s)$$

$$\{f_o\} = \{f_i\} \times \{h\}_A$$

$$\{f_i\} = \Delta t, \quad 2\Delta t, \quad 3\Delta t, \quad 4\Delta t, \quad 5\Delta t, \quad 6\Delta t, \dots$$

$$\{h\}_A = 1/\Delta t, \quad -1/\Delta t,$$

---


$$\begin{array}{cccccc} 1, & 2, & 3, & 4, & 5, & 6, \dots \\ -1, & -2, & -3, & -4, & -5, & \dots \end{array}$$


---

$$\{f_o\} = 1, \quad 1, \quad 1, \quad 1, \quad 1, \dots$$

(iv) Suppose we now try to differentiate the step. We get the unit impulse as we should

$$\{f_i\} = 1, \quad 1, \quad 1, \quad 1, \quad 1, \quad 1, \dots$$

$$\{h\}_A = 1/\Delta t, \quad -1/\Delta t,$$


---

	$1/\Delta t,$	$1/\Delta t,$	$1/\Delta t,$	$1/\Delta t,$	$1/\Delta t,$	$1/\Delta t, \dots$
	$-1/\Delta t,$	$-1/\Delta t,$	$-1/\Delta t,$	$-1/\Delta t,$	$-1/\Delta t,$	$\dots$
$\{f_o\}$	$= 1/\Delta t,$	$0,$	$0,$	$0,$	$0,$	$0, \dots$
(v) Finally we differentiate a sine wave, $\Delta t = 1/10$ .						
$\{f_i\}$	$=$	$0.1,$	$0.199,$	$0.296,$	$0.389,$	$0.479, 0.564, 0.644, 0.717$
$\{h\}_A$	$=$	$10,$	$-10$			
		$1,$	$1.99,$	$2.96,$	$3.89,$	$4.79, 5.64, 6.44, 7.17$
		$-1,$	$-1.99,$	$-2.96,$	$-3.89,$	$-4.79, -5.64, -6.44$
$\{f_o\}$	$=$	$1,$	$0.99,$	$0.97,$	$0.93,$	$0.90, 0.85, 0.80, 0.73$
$\{f_i\}$	(cont'd.)	$0.783,$	$0.841,$	$0.891,$	$0.932,$	$0.964, 0.985, 0.997$
		$7.83,$	$8.41,$	$8.91,$	$9.32,$	$9.64, 9.85, 9.97$
	(cont'd.)	$-7.17,$	$-7.83,$	$-8.41,$	$-8.91,$	$-9.32, -9.64, -9.85$
$\{f_o\}$	(cont'd.)	$0.66,$	$0.57,$	$0.50,$	$0.41,$	$0.32, 0.21, 0.12$

It can be recognized that  $\{f_o\}$  is a cosine wave that is the derivative of the input.

Let us now take the problem that is inverse to the previous one, that is, to specify an output that is the integral of the input. Here again, to simplify the arithmetic, we choose without any loss of generality an input that is the unit step function and an output that is the unit ramp.

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}} = \frac{\Delta t, 2\Delta t, 3\Delta t, 4\Delta t, \dots}{1, 1, 1, 1, \dots}$$

$$1, 1, 1, 1, \dots \left| \begin{array}{cccccc} \Delta t, & \Delta t, & \Delta t, & \Delta t, & \Delta t, & \dots \\ \Delta t, & 2\Delta t, & 3\Delta t, & 4\Delta t, & 5\Delta t, & \dots \\ \Delta t, & \Delta t, & \Delta t, & \Delta t, & \Delta t, & \dots \\ \hline & \Delta t, & 2\Delta t, & 3\Delta t, & 4\Delta t, & \dots \\ & \Delta t, & \Delta t, & \Delta t, & \Delta t, & \dots \\ \hline & & \Delta t, & 2\Delta t, & 3\Delta t, & \dots \\ & & \Delta t, & \Delta t, & \Delta t, & \dots \\ \hline & & & \Delta t, & 2\Delta t, & \dots \\ & & & \Delta t, & \Delta t, & \dots \\ \hline & & & & \Delta t, & \dots \\ & & & & \Delta t, & \dots \end{array} \right.$$



As we know, the impulse response is the unit step and the system function is  $1/s$ . Convolution of any function with the unit step time operator results in integration of the input function, as can be seen by the following example of the cosine function (with  $\Delta t = 0.1$ ):

$$\begin{array}{r}
 \{f_i\} \quad 1, 0.99, 0.97, 0.93, 0.90, 0.85, 0.80, 0.73, \dots \\
 \{h\}_A \quad 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, \dots \\
 \hline
 \quad 0.1, 0.099, 0.097, 0.093, 0.090, 0.085, 0.080, 0.073, \dots \\
 \quad \quad 0.1, 0.099, 0.097, 0.093, 0.090, 0.085, 0.080, \dots \\
 \quad \quad \quad 0.1, 0.099, 0.097, 0.093, 0.090, 0.085, \dots \\
 \quad \quad \quad \quad 0.1, 0.099, 0.097, 0.093, 0.090, \dots \\
 \quad \quad \quad \quad \quad 0.1, 0.099, 0.097, 0.093, \dots \\
 \quad \quad \quad \quad \quad \quad 0.1, 0.099, 0.097, \dots \\
 \quad \quad \quad \quad \quad \quad \quad 0.1, 0.099, \dots \\
 \quad \quad \quad \quad \quad \quad \quad \quad 0.1, \dots \\
 \hline
 \{f_o\} \quad 0.1, 0.199, 0.296, 0.389, 0.479, 0.564, 0.644, 0.717, \dots
 \end{array}$$

which checks the sine function.

(vii) We take next a case where we feed a sine wave into a Zobel constant - k type low-pass filter and get a

$$\begin{aligned}
 \{f_o\} &= \{ 0.0138, 0.0621, 0.1388, 0.2267, 0.3048, 0.3516, 0.3525, \\
 &\quad 0.3001, 0.1997, 0.0657, -0.0803, -0.2131, -0.3099, -0.3540, \\
 &\quad -0.3369, -0.262, -0.1417, 0.0030, 0.1475, 0.2664, 0.3391, \dots \} \\
 \{f_i\} &= \{ 0.406, 0.743, 0.95, 0.994, 0.866, 0.588, 0.208, -0.208, -0.588, \\
 &\quad -0.866, -0.994, -0.95, -0.743, -0.406, 0, 0.406, 0.743, \\
 &\quad 0.95, 0.994, 0.866, 0.588, 0.208, -0.208, -0.588, \dots \}
 \end{aligned}$$

Then

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}} = \{ 0.034, 0.091, 0.096, 0.087, 0.072, 0.053, \\
 \quad 0.036, 0.020, 0.009, 0.002, -0.001, -0.002, \\
 \quad -0.002, -0.002, -0.001, 0, 0, 0, \dots \}$$

which is the impulse response of the filter.

Let us consider some discontinuous inputs and outputs, since these are frequently of interest in time-domain synthesis of networks.

(viii) Let  $f_i(t)$  be a rectangular pulse with discontinuities at  $t = 0$  and  $t = 0.5$ , and  $f_o(t)$  be a triangular pulse with discontinuities in the slope at  $t = 0$ ,  $t = 0.5$  and  $t = 1$ .

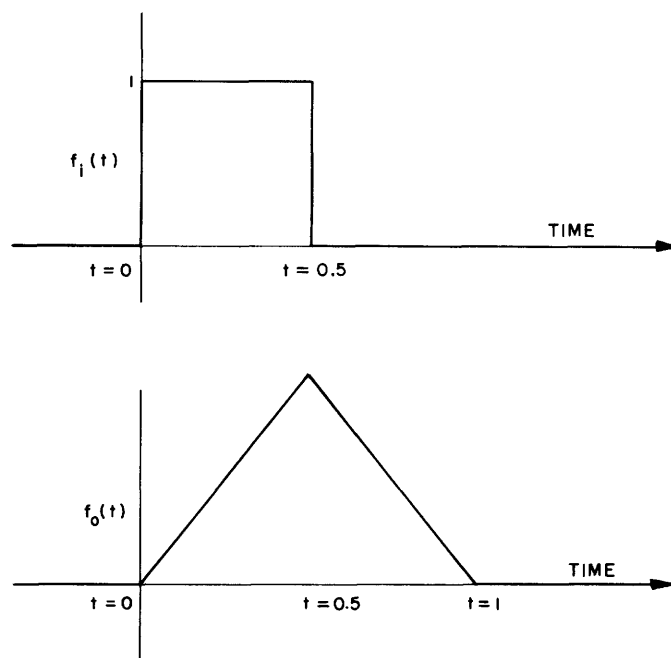


Fig. 7

Then, for  $\Delta t = 0.1$ ,

$$\{f_i\} = \{1, 1, 1, 1, 1, 0, 0, 0, 0, \dots\}$$

$$\{f_o\} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1, 0, 0, \dots\}$$

$$\begin{array}{r}
 1, 1, 1, 1, 1, 0, 0, \dots \quad \begin{array}{l}
 \begin{array}{r}
 0.1, 0.1, 0.1, 0.1, 0.1, \quad 0, \quad 0, \quad 0, \quad 0, 0, 0, \dots \\
 \hline
 0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1, 0, 0, 0, \dots \\
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 \hline
 0.1, 0.2, 0.3, 0.4, 0.4, \\
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 \hline
 0.1, 0.2, 0.3, 0.3, 0.3, \\
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 \hline
 0.1, 0.2, 0.2, 0.2, 0.2, \\
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 \hline
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 0.1, 0.1, 0.1, 0.1, 0.1, \\
 \hline
 0, \quad 0, \quad 0, \quad 0, \quad 0, 0, 0, 0,
 \end{array}
 \end{array}
 \end{array}$$

A rectangular impulse response with height 1 and discontinuities at  $t = 0$  and  $t = 0.5$  is represented by  $\{h\}_A$ .

(ix) If the input is to be an exponential pulse and the network is to give the same

triangular output, then the impulse response for  $\Delta t = 0.1$  will be

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}} = \frac{\{0.1, 0.2, 0.3, 0.4, 0.5, 0.4, 0.3, 0.2, 0.1, 0, 0, 0, 0, 0, \dots\}}{\{0.9048, 0.8187, 0.7408, 0.6703, 0.6065, 0.5488, 0.4966, 0.4493, 0.4066, \dots\}}$$

and carrying through the synthetic division we obtain

$$\{h\}_A = \{0.11, 0.121, 0.1315, 0.14207, 0.1526, -0.0579, -0.06847, -0.07897, -0.91, -0.9865, 0, 0, 0, \dots\}$$

(x) To conclude this set of examples, let the input be the same as above but suppose the output to be  $4t \exp(-4t)$  with a discontinuity at  $t = 1$  when a delayed  $-4t \exp(-4t)$  is to be superimposed, that is,

$$\{f_o\} = \{0.27, 0.36, 0.361, 0.323, 0.270, 0.217, 0.170, 0.130, 0.098, 0.073, -0.214, -0.320, -0.333, -0.302, -0.255, -0.208, -0.163, -0.125, -0.094, -0.071, -0.054, -0.039, -0.028, -0.021, -0.015, \dots\}$$

as in Fig. 8 with  $\Delta t = 0.1$ .

Then

$$\{h\}_A = \frac{\{f_o\}}{\{f_i\}} = \{0.40, 0.268, 0.18, 0.12, 0.08, 0.054, 0.036, 0.024, 0.016, 0.011, -0.40, -0.268, -0.18, -0.12, -0.08, -0.054, -0.036, -0.024, -0.016, -0.011, 0, 0, \dots\}$$

which is seen to be an exponential impulse response  $4 \exp(-4t)$  followed by a negative one with a delay of one time unit. We can check this result from

$$\frac{F_o(s)}{F_i(s)} = \frac{4/(s+4)^2 - [4 \exp(-s)]/(s+4)^2}{1/s + 4} = \frac{4}{s+4} - \frac{4 \exp(-s)}{s+4}$$

### 3.4 THE CONVERGENCE OF THE PROCEDURE FOR OBTAINING THE IMPULSE RESPONSE

We close this section with a few observations concerning the convergent nature of this synthetic division method from the viewpoint of the choice of  $\Delta t$ . We may think of this choice of interval as the rate at which we sample the time function data. Obviously, if we took our samples very widely spaced, we would be losing the essential information contained in the data. For that reason we need to take the  $\Delta t$  sufficiently close to ensure that any oscillations in either  $f_i(t)$  or  $f_o(t)$  are adequately brought out in the time sequences.

If we had chosen  $\Delta t$  so that the above limit is taken care of, what happens to the  $h(t)$  that we would get if we began making  $\Delta t$  smaller; in other words, if we sample more frequently? We would expect that the successive  $h(t)$  which we may signify by  $h_1(t)$ ,  $h_2(t)$ ,  $h_3(t)$  corresponding to  $\Delta t_1$ ,  $\Delta t_2$ ,  $\Delta t_3$ , and so on, would approach more and more closely to a limiting form which, in the case of  $f_i(t)$  and  $f_o(t)$  (which are known analytic functions),

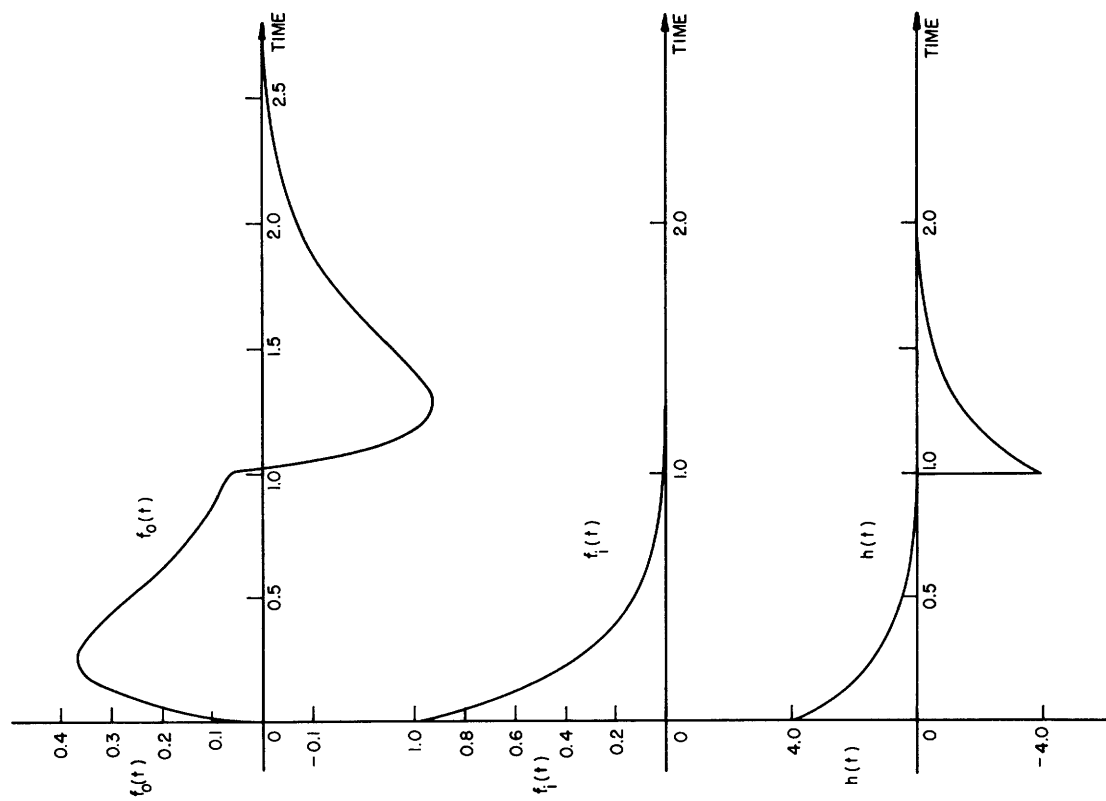


Fig. 8

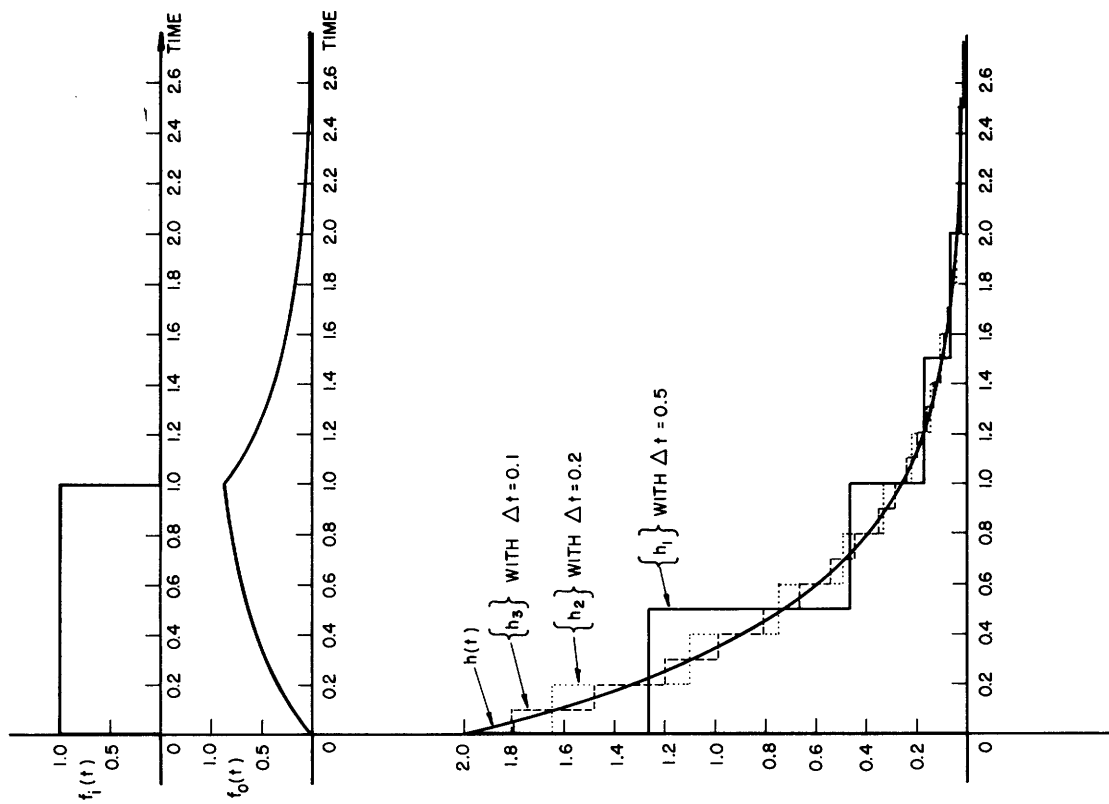


Fig. 9

would be some definite curve. We will illustrate by a few examples that this is, indeed, the case.

a. Let  $f_i(t)$  be a rectangular pulse with discontinuities at  $t = 0$  and  $t = 1$ , and let  $f_o(t) = 1 - \exp(-2t)$  with a delayed  $-[1 - \exp(-2t)]$  starting at  $t = 1$ , so that the input and output are as shown in Fig. 9.

We will first use  $\Delta t_1 = 0.5$ . From this we get

$$\begin{aligned} \{h_1\}_A &= \frac{\{f_o\}}{\{f_i\}} = \frac{\{0.6321, 0.8647, 0.3181, 0.1170, 0.043, 0.0158\}}{\{1, 1, 0, 0, 0, 0, \dots\}} \\ &= \{0.6321, 0.2326, 0.0855, 0.0315, 0.0115, 0.0043\} \end{aligned}$$

Next, we use  $\Delta t_2 = 0.2$  and get for

$$\begin{aligned} \{f_o\} &= \{0.33, 0.55, 0.70, 0.79, 0.865, 0.58, 0.388, 0.26, 0.174, 0.177, 0.0785, \\ &\quad 0.053, 0.035, 0.024, 0.016, \dots\} \\ \{f_i\} &= \{1, 1, 1, 1, 1, 0, 0, 0, \dots\} \end{aligned}$$

and

$$\{h_2\}_A = \frac{\{f_o\}}{\{f_i\}} = \{0.33, 0.22, 0.15, 0.098, 0.067, 0.045, 0.029, 0.022, \dots\}$$

which yields a  $\{h_2\}$  that is seen to approximate  $h(t) = 2 \exp(-2t)$  more closely than  $\{h_1\}$ .

If we do the same with  $\Delta t_3 = 0.1$ , we obtain

$$\{h_3\}_A = \{0.18, 0.148, 0.12, 0.099, 0.0814, 0.0667, 0.0546, 0.044, 0.035, 0.028, 0.024, 0.019, 0.015, \dots\}$$

which continues the convergence on  $h(t)$ .

Note that all three approximations  $\{h_1\}_A$ ,  $\{h_2\}_A$ , and  $\{h_3\}_A$  have the common property that the areas are correct to within the tolerances that are set by our choice of ordinate values of  $f_o(t)$  and  $f_i(t)$ . That this must be so is made clear by a consideration of the reverse process; namely the convolution of  $f_i(t)$  with  $h(t)$  to obtain  $f_o(t)$ , which involves an integration, that is, it is the areas under the  $h(t)$  curve which are really instrumental in producing  $f_o(t)$  from  $f_i(t)$ . We can see by section 3.1 that we may replace convolution by  $\{f_i\} \times \{h\}_A$  to obtain  $\{f_o\}$ . In order to obtain the values of  $\{f_o\}$ , picked from  $f_o(t)$  corresponding to any particular  $\Delta t$  with an analogous  $\{f_i\}$ , we obviously must have  $\{h\}_A$  equal to the areas under the true  $h(t)$  curve, even though the shape of the approximating curves may be different and discontinuous.

b. Let us see whether this convergent behavior is maintained if the true impulse response itself is a discontinuous curve. For this purpose let us choose  $f_i(t) = \exp(-t)$  and  $f_o(t)$  a triangular pulse with discontinuities in the slope at  $t = 0, 0.5$  and  $1$ . The curves in Fig. 10 illustrate that the process of choosing  $\Delta t$  successively smaller does lead nearer and nearer to the true impulse response.

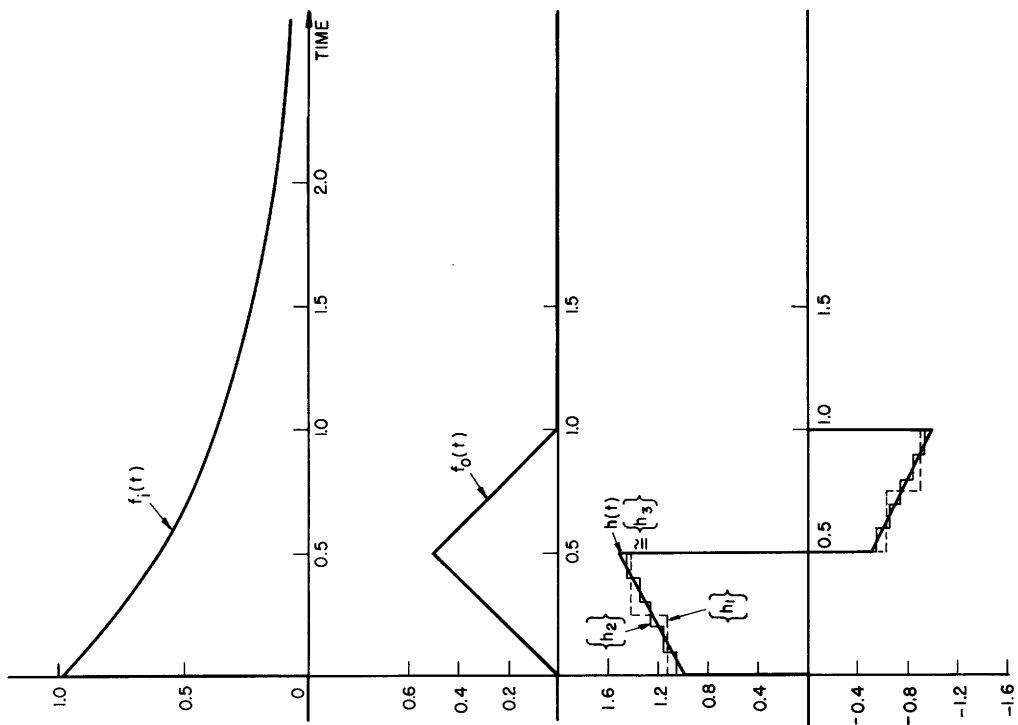


Fig. 10

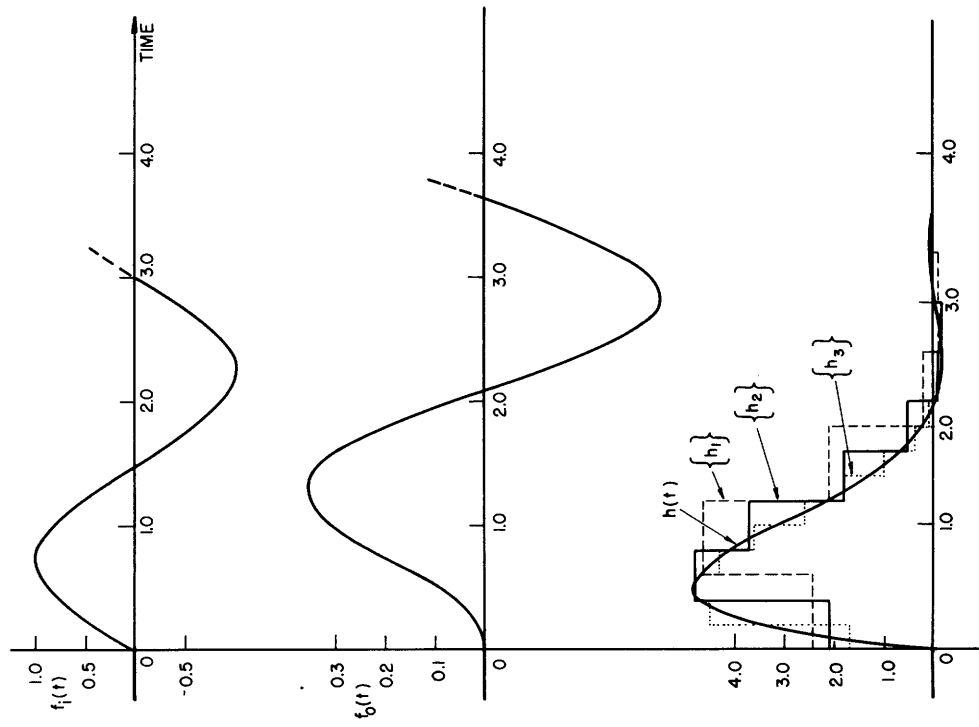


Fig. 11

$$\{h_1\}_A = \frac{\{f_o\}}{\{f_i\}} \text{ for } \Delta t_1 = 0.25 = \frac{\{0.25, 0.5, 0.25, 0, 0, 0, \dots\}}{\{0.88, 0.69, 0.53, 0.41, 0.32, 0.25, 0.19, \dots\}}$$

$$\begin{array}{r} 0.283, 0.357, -0.159, -0.222, 0, 0, 0, \\ 0.88, 0.69, 0.53, 0.41, 0.32, 0.25, 0.19, \dots \left| \begin{array}{l} 0.25, 0.5, 0.25, 0, 0, 0, 0 \\ 0.25, 0.186, 0.15, 0.116, 0.09, 0.07, 0.054 \\ 0.314, 0.10, -0.116, -0.09, -0.07, -0.054 \\ 0.314, 0.24, 0.19, 0.146, 0.114, 0.09 \\ -0.14, -0.306, -0.236, -0.184, -0.144 \\ -0.14, -0.11, -0.084, -0.065, -0.057 \\ -0.196, -0.152, -0.119, -0.093 \\ -0.196, -0.152, -0.119, -0.093 \\ 0, 0, 0, 0 \end{array} \right. \end{array}$$

$\{h_2\}$  and  $\{h_3\}$  are obtained by using  $\Delta t_2 = 0.1$  and  $\Delta t_3 = 0.01$ .

c. Finally, we give the case of a sinusoidal input  $f_i(t) = \sin 20t$  passing through a filter to yield an output shown by Fig. 11. We first take  $\Delta t_1 = 0.06$ , which gives a rather coarse approximation to the impulse response since the choice of interval is pretty near the limit. The successive approximations with  $\Delta t_2 = 0.04$  and  $\Delta t_3 = 0.02$  are fair and good.

$$\{h_1\}_A = \frac{\{0.139, 0.351, 0.199, -0.213, -0.337, 0.003, \dots\}}{\{0.95, 0.588, -0.588, -0.95, 0, 0.95, \dots\}}$$

## IV. CRITERIA FOR PHYSICAL REALIZABILITY

### 4.1 THE CRITERIA ON INPUT-OUTPUT PAIRS THAT CORRESPOND TO PHYSICALLY REALIZABLE NETWORKS

We have shown a procedure for obtaining the impulse response from any pair of input and output given either as functions of time, or as a pair of curves, or simply as sequences of ordinate values at specified time intervals. This brings us a big step closer to the synthesis of the network for achieving this input-output relationship. But the very mention of the word "synthesis" makes us question the physical realizability of the network. This is a very natural question because if the network is not going to be physically realizable, there is no practical point in bothering to carry through the design calculations for it. Therefore, we need to have some simple criteria that we can apply in order to be able to say at the outset whether or not the network is going to be physically realizable.

At this juncture, let us briefly remark that the networks we are interested in are linear, passive networks and consist of only real, positive resistances, inductances, capacitances, and some kind of transformers. This means that the term "physically realizable" as used here refers to the fact that we can design a network using only these elements as its components connected in some physical configuration. Hence we say in this context that a network is not physically realizable if the synthesis calls for negative resistances (or conductances) or, in general, if the impedances or admittances require negative real parts. This means that, in the language of theory of functions of a complex variable, we are not allowed network system functions  $H(s)$  which have poles in the right half of the complex frequency plane.

Let us now translate this statement, by the inverse Laplace transform, into a statement regarding the time functions. Since  $f_1(t)$  is the excitation, and the network is supposed to be passive, obviously the response cannot start before the excitation is applied. Therefore the first thing we should look for in a pair of input-output time functions is to see if the output is nonzero before  $t = 0$ , which is defined as the instant at which the excitation  $f_1(t)$  is applied to the network.

Secondly, since a pole of  $H(s)$  in the right half  $s$ -plane gives rise to a time function that increases exponentially with time, any input-output pair which yields an impulse response  $h(t)$  of that positive exponential type will indicate that we cannot realize it by a passive network.

In Figs. 12(a) and 12(b) there is a catalogue of some of the possible impulse responses. Figures 13(a) and 13(b) show the types that are not physically realizable. The two criteria above are valid in general and are easily evident by an inspection of the waveform of the time functions. There is no difficulty in seeing whether or not the first criterion (response must not start before excitation is applied) is violated. If it is violated, it will be apparent as shown in Fig. 13(a) and we may straightway stop considering it for any synthesis procedures.



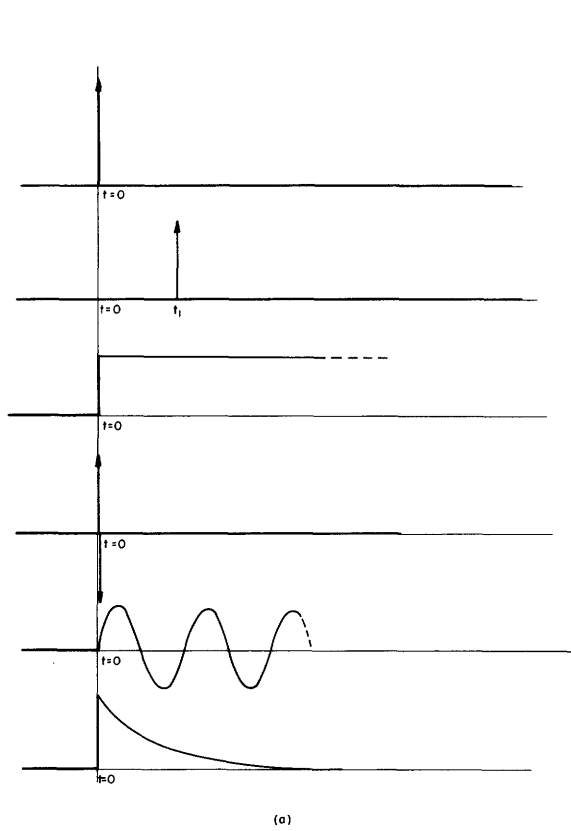


Fig. 12(a)

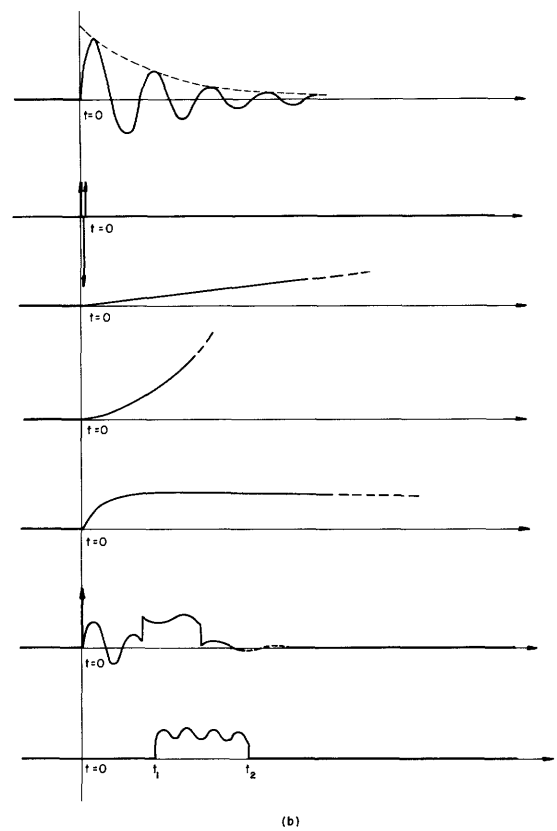


Fig. 12(b)

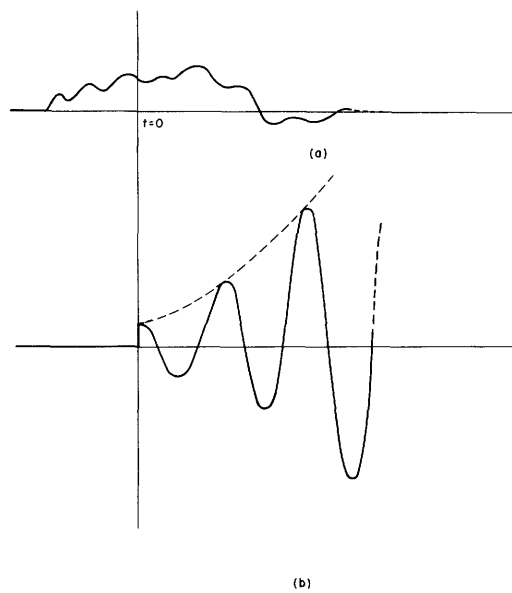


Fig. 13

If an input-output pair passes this first test, then we go ahead and get the impulse response. We then take a look at this  $h(t)$  and see if it passes the second criterion. To those who have done network synthesis in the frequency domain starting with  $Z(s) = P(s)/Q(s)$ , the above two tests are the analogs of the two tests we performed on  $Q(s)$  to see if  $Z(s)$  is a physically realizable transfer function, namely (1) does  $Q(s)$  have any complex or negative real or missing coefficients? This is evident by inspection. If it does, reject it as it is not a Hurwitz polynomial. Thus  $Z(s)$  cannot be realized. (2) After  $Q(s)$  passes the above test, carry out the division procedure known as the Hurwitz test to see if  $Q(s)$  has any poles in the right half  $s$ -plane. Computationally this is similar to our process for finding  $\{h\}$ .

We may say a little more about the input and output functions. Note that earlier in section 2.2,  $f_i(t)$  and  $f_o(t)$  must be single-valued, real, bounded functions of time. These restrictions are no hardship at all, since in any practical situation all the inputs and outputs will surely fulfill these simple conditions. Furthermore, since we are not responsible for generating  $f_i(t)$ , it may contain exponentially increasing components provided that it does not continue indefinitely, because then it would no longer satisfy the boundedness limitation

$$\int_0^{\infty} |f_i(t)| dt < \infty$$

However, this again is not a drastic requirement, because sooner or later the input will be switched off. If the input does contain positive exponentials, then the output may also contain the same without the impulse response necessarily having such positive exponentials. For example, if we consider a device such as a voltage divider, it is certainly realizable with real resistances and will produce an exponentially increasing output with an exponentially increasing input. This simple case was cited to warn against hasty conclusions after looking only at the output. We must always consider the input and output as an inseparable pair and remember that it is the impulse response that actually determines whether or not the network is realizable.

We might also add, for those who are used to thinking in terms of zeros and poles of transfer functions, that in

$$H(s) = \frac{F_o(s)}{F_i(s)}$$

if both  $F_i(s)$  and  $F_o(s)$  contain the same pole in the right half-plane, that pole will cancel out in  $H(s)$ ; then  $H(s)$  will be realizable. Moreover, even if  $H(s)$  does contain a right half-plane pole but  $h(t)$  is forced to be of finite duration, then that pole will be cancelled by right-hand plane zeros at the same point. For example, let  $h(t) = \exp(at)$ , for  $0 < t < t_1 = 0$ , for all other  $t$ . This gives

$$H(s) = \frac{1}{s-a} - \frac{\exp(at_1) \exp(-t_1 s)}{s-a} = \frac{1 - \exp[-t_1(s-a)]}{(s-a)}$$

$$= \frac{1}{(s-a)} \left[ 1 - 1 + (s-a)t_1 - (s-a)^2 \frac{t_1^2}{2} + \frac{(s-a)^3 t_1^3}{3!} - \dots \right] = t_1 - (s-a) \frac{t_1^2}{2!} + (s-a)^2 \frac{t_1^3}{3!} - \dots$$

which has no poles in the right half-plane, only zeros. These are allowed for transfer functions.

But the case where  $F_i(s)$  has zeros in the right half-plane is not so easily dismissed, because if  $F_o(s)$  does not contain the same zero, then  $H(s)$  will now have a right half-plane pole that may not be cancelled unless we modify either the input or the output. Because this case is interesting, we will consider it in greater detail in section 4.2.

#### 4.2 THE SPECIAL CASE OF HIDDEN ZEROS OF THE TRANSFORM OF THE INPUT FUNCTION IN THE RIGHT HALF S-PLANE

We use the term "hidden zeros" because, by merely looking at the curve of the input time function, it is frequently hard to tell if its transform contains zeros in the right half-plane. This is in contrast to poles in the right half-plane that manifest themselves prominently in the time function curve. The curves of Fig. 14 illustrate some input time functions (with such hidden zeros).

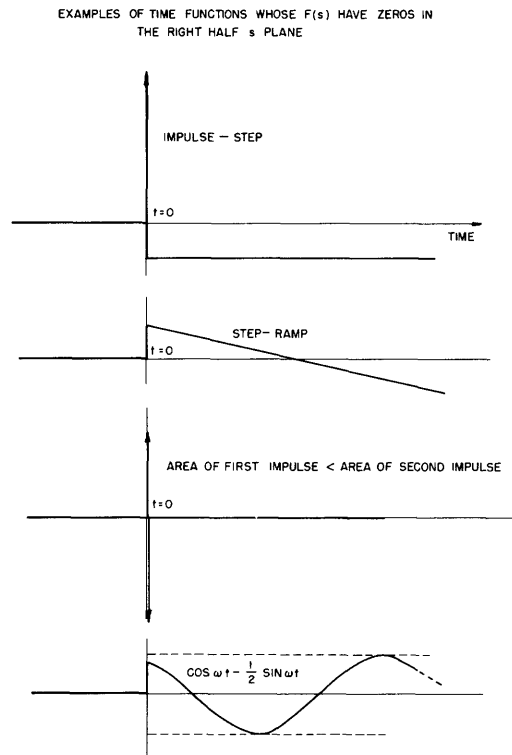


Fig. 14

In the case of the sinusoidal functions whose zeros in the right half-plane are particularly difficult to notice, we can dispose of that zero by adding sufficient dc bias to  $f_i(t)$ . The reason is made clear when we consider the Laplace transforms.

For  $f_i(t) = \exp(-at)(\cos \beta t - a_o \sin \beta t)$ , where  $a_o \beta > a$  we get

$$F_i(s) = \frac{s - (a_o \beta - a)}{s^2 + 2as + a^2 + \beta^2}$$

which has a zero in the right half-plane, since  $a_o \beta > a$ .

Now add  $k/s$ , which corresponds to a positive dc bias of  $k$ .

$$\begin{aligned} F_i^*(s) &= \frac{k}{s} + \frac{s - (a_o \beta - a)}{s^2 + 2as + a^2 + \beta^2} = \frac{ks^2 + k2as + ka^2 + k\beta^2 + s^2 - s(a_o \beta - a)}{s(s^2 + 2as + a^2 + \beta^2)} \\ &= \frac{s^2(1+k) + s(2ak - a_o \beta + a) + (ka^2 + k\beta^2)}{s(s^2 + 2as + a^2 + \beta^2)} \end{aligned}$$

The roots of the numerator are

$$\frac{-(2ak + a - a_o \beta) \pm \left[ (2ak + a - a_o \beta)^2 - 4(1+k)(ka^2 + k\beta^2) \right]^{1/2}}{2(1+k)}$$

If  $\left[ (2ak + a - a_o \beta)^2 - 4(1+k)(ka^2 + k\beta^2) \right]^{1/2}$  is real, it can never exceed  $(2ak + a - a_o \beta)$  because  $4(1+k)(ka^2 + k\beta^2) > 0$ ; hence the roots will always be in the left half-plane, if  $k$  is chosen large enough to make  $2ak + a - a_o \beta > 0$ . Thus  $F_i^*(s)$  now has no more zeros in the right half  $s$ -plane, while  $f_i^*(t)$  still has the same waveform as  $f_i(t)$ , except for a shift in bias level.

#### 4.3 SEPARATION OF PULSE SIGNALS OCCURRING SIMULTANEOUSLY IN THE INPUT

We now take up a topic that is of theoretical significance in time-domain work, namely, the separation of two signals whose frequency bands overlap and, therefore, cannot be separated by bandpass filter techniques, but whose time functions are distinguishable enough to offer the hope that they might be separated using time-domain synthesis methods.

Suppose that the input consists of a rectangular pulse of duration  $t = 1$  on which are superimposed a number of smaller pulses of shorter duration, say  $t = 0.2$ . For example, let

$$\{f_i\} = \{1, 1, 1, 1+b, 1+b, 1, 1+b, 1+b, 1, 1, 0, \dots\}$$

where  $b$  = the heights of the smaller pulses. The problem then is to recover only the large pulse to within a tolerance of  $\pm\mu$ .

$$\begin{aligned}\{h\}_A &= \frac{\{1, 1, 1, 1 + \mu, 1 + \mu, 1, 1 + \mu, 1 + \mu, 1, 1, 0, \dots\}}{\{1, 1, 1, 1 + b, 1 + b, 1, 1 + b, 1 + b, 1, 1, 0, \dots\}} \\ &= \{1, 0, 0, (\mu - b), 0, -(\mu - b), (\mu - b)(1 - b), 0, (\mu - b)(2b - 1), \dots\}\end{aligned}$$

If now we send in an input that consists of the long duration pulse only, then the output of the above network would be

$$\begin{aligned}\{f_o\}_1 &= \{f_i\}_1 \times \{h\}_A \\ &= \{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, \dots\} \\ &\quad \times \{1, 0, 0, (\mu - b), 0, -(\mu - b), (\mu - b)(1 - b), 0, (\mu - b)(2b - 1), \dots\} \\ &= \{1, 1, 1, 1 + \mu - b, 1 + \mu - b, 1, 1 + \mu - \mu b - b + b^2, \\ &\quad 1 + \mu - \mu b + b^2, 1 + \mu b - b^2, 1 + \mu b - b^2, \dots\}\end{aligned}$$

For this  $\{f_o\}_1$  to approximate  $\{1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, \dots\}$  to within the tolerance  $\pm\mu$ , the condition on  $b$  will be  $1 + \mu - b \geq 1 - \mu$ , or  $b \leq 2\mu$ . This satisfies the requirements on  $1 + \mu > 1 + \mu - \mu b - b + b^2 > 1 - \mu$  and  $1 + \mu > 1 + \mu b - b^2 > 1 - \mu$  also. Thus, if the interfering signal is not greater than twice the allowable tolerance, it seems that separation of the two signals is possible to recover the longer duration pulse only.

## V. METHODS FOR OBTAINING THE SYSTEM FUNCTION

### 5.1 METHOD OF CALCULATING $H(s)$ , THE SYSTEM FUNCTION, ALGEBRAICALLY FROM THE IMPULSE RESPONSE $h(t)$ , USING THE POWER SERIES EXPANSION

We now come to the main problem of time-domain synthesis. After the primary problem of obtaining the impulse response from the specified input and output (sec. III) has been solved, and after this impulse response has passed the criteria for physical realizability (sec. IV), how are we to get  $H(s)$ , the system function, from this impulse response? A ready answer may be given if the impulse response is known as an analytic function of time,  $h(t)$ ; because, in this case, (which is, however, very rare in practical engineering) we need only to find

$$H(s) = \int_0^{\infty} h(t) \exp(-st) dt \quad (25)$$

or more easily, look up a table of Laplace transforms. Of course, if this transform turns out to be a transcendental function, our troubles will not yet be over. We shall see later how to get a rational function approximation.

However, we usually do not know what  $h(t)$  is, since to get the impulse response as a function of time means that we have solved the integral equation obtained from Eq. 12, which, in general, is not possible. Therefore, we have only the impulse response available to us as a time sequence or as a curve (discontinuous approximation or a smooth curve obtained from it). The Dirichlet series representation for the Laplace-Stieltjes integral again comes to our aid. We have seen in section 3.4 that the method of synthetic division yields a time sequence (or the corresponding curve with time as the abscissa) in which the areas during the intervals  $\Delta t$  are correct to those of the true impulse response. The coefficients of the Dirichlet series, Eq. 22, correspond to the areas under the curve in the chosen intervals. Therefore, we may obtain the Dirichlet series representation of  $H(s)$  directly from the curve or area time sequence  $\{h\}_A$ .

However, we are not satisfied with this form of  $H(s)$  because it does not easily lend itself to frequency-domain synthesis. We would like it in rational function form, that is, as a ratio of finite polynomials in  $s$ , because then there exist a number of procedures by which we can deal with this rational function and get the physical network. Therefore, from the point of view of the time-domain synthesist, he may consider his part of the work as accomplished when he can write  $H(s) = P(s)/Q(s)$  for any input-output pair in the time domain.

Three original methods are presented in this report for calculating  $H(s)$  as a rational function. They are based on two totally different theoretical approaches. All of them have the common virtue (very precious to an engineer) of being mathematically simple to carry out, because actual computations involve only straightforward algebra of an elementary nature; namely, the solution of a finite set of linear, simultaneous, algebraic

equations in the same number of real variables. These are the unknown coefficients of the polynomials  $P(s)$  and  $Q(s)$  and are guaranteed to be real because of the physical realizability of the network.

We shall now give the method that is based on a power series expansion valid for the whole finite  $s$ -plane.

We have

$$\exp(-\tau_v s) = 1 - \frac{(\tau_v s)}{1!} + \frac{(\tau_v s)^2}{2!} - \frac{(\tau_v s)^3}{3!} + \frac{(\tau_v s)^4}{4!} \dots (\tau_v s)^2 < \infty \quad (26)$$

and so we may write for the Dirichlet representation of the Laplace transform

$$H(s) = \sum_{v=1}^{\mu} a_v \exp(-\tau_v s) \quad (27)$$

the power series

$$\begin{aligned} H(s) = & a_1 \left[ 1 - \tau_1 s + \frac{(\tau_1 s)^2}{2!} - \frac{(\tau_1 s)^3}{3!} + \frac{(\tau_1 s)^4}{4!} \dots \right] \\ & + a_2 \left[ 1 - \tau_2 s + \frac{(\tau_2 s)^2}{2!} - \frac{(\tau_2 s)^3}{3!} + \frac{(\tau_2 s)^4}{4!} \dots \right] \\ & + \dots \dots \dots \\ & + a_{\mu} \left[ 1 - \tau_{\mu} s + \frac{(\tau_{\mu} s)^2}{2!} - \frac{(\tau_{\mu} s)^3}{3!} + \frac{(\tau_{\mu} s)^4}{4!} \dots \right] \end{aligned} \quad (28)$$

or

$$\begin{aligned} H(s) = & \sum_{v=1}^{\mu} a_v - s \sum_{v=1}^{\mu} a_v \tau_v + \frac{s^2}{2!} \sum_{v=1}^{\mu} a_v \tau_v^2 - \frac{s^3}{3!} \sum_{v=1}^{\mu} a_v \tau_v^3 + \dots \\ = & b_0 + b_1 s + b_2 s^2 + b_3 s^3 + \dots \end{aligned} \quad (29)$$

We want to have this system function  $H(s)$  as a ratio of two finite polynomials. So we merely write

$$H^*(s) = \frac{P(s)}{Q(s)} = \frac{p_0 + p_1 s + p_2 s^2 + \dots + p_n s^n}{q_0 + q_1 s + q_2 s^2 + \dots + q_m s^m}$$

and equate this to the series for  $H(s)$  in Eq. 29. Thus we get

$$\frac{p_0 + p_1 s + p_2 s^2 + \dots + p_n s^n}{q_0 + q_1 s + q_2 s^2 + \dots + q_m s^m} = b_0 + b_1 s + b_2 s^2 + b_3 s^3 + \dots$$

or

$$p_0 + p_1 s + \dots + p_n s^n = (q_0 + q_1 s + q_2 s^2 + \dots + q_m s^m) \times (b_0 + b_1 s + b_2 s^2 + \dots) \quad (30)$$

The following set of equations results by equating coefficients of like powers of  $s$ , in order for  $H^*(s)$  to be equal to  $H(s)$

$$\begin{aligned} p_0 &= b_0 q_0 \\ p_1 &= b_1 q_0 + b_0 q_1 \\ p_2 &= b_2 q_0 + b_1 q_1 + b_0 q_2 \\ &\dots \dots \dots \\ p_n &= b_n q_0 + b_{n-1} q_1 + b_{n-2} q_2 + \dots + b_0 q_n \\ 0 &= b_{n+1} q_0 + b_n q_1 + b_{n-1} q_2 + \dots + b_1 q_n + b_0 q_{n+1} \end{aligned} \quad (31)$$

and so on.

We assume here that  $m > n$  because if  $H(s)$  originally had its numerator of higher degree, we could always get it into proper fraction form by division. The terms corresponding to constants and powers of  $s$  will appear very prominently in the time function as impulses and doublets and triplets and so on, and therefore may be easily recognized. Thus, we need concern ourselves only with that part of  $H(s)$  that corresponds to the impulse response minus all these singularity functions. That part of  $H(s)$  will have a rational function approximation which has its denominator of higher degree than the numerator. More will be said about this relative degree of numerator and denominator later, but let us go back to the preliminary description of this first method.

There are an infinite number of equations in Eq. 31, but we take only the first  $m + n + 2$  equations to determine the  $m + n + 2$  unknowns  $p_0, p_1, p_2, \dots, p_n, q_0, q_1, q_2, \dots, q_m$ . This ensures that we have matched the first  $m + n + 2$  coefficients of the power series corresponding to  $H(s)$  with that obtained by division of  $P(s)$  by  $Q(s)$ , because we actually used the coefficients  $b_0, b_1, b_2, \dots, b_{m+n+1}$  of the power series. But what about the later coefficients; and what errors do we commit by not taking them into account?

In the first case, if  $H(s)$  is itself a rational function with numerator polynomial of degree  $N$  and denominator polynomial of degree  $M$ , then it has been found that by choosing  $n = N$  and  $m = M$ , where  $n$  and  $m$  are the degree of the polynomials  $P(s)$  and  $Q(s)$  of  $H^*(s)$ , and by using just the  $m + n + 2$  equations necessary to determine the unknown coefficients  $p_0, p_1, p_n, q_0, q_1, \dots, q_m$ , we get these precisely equal to the



coefficients of the polynomials in the original  $H(s)$ . This is not surprising when we reflect that since the power series, Eq. 29, would, in the case of a rational function  $H(s)$ , be obtained by dividing the numerator polynomial by the denominator polynomial, and if we match this quotient with that obtained from  $H^*(s) = P(s)/Q(s)$ , then although we actually match only the beginning terms of the power series, all subsequent terms of the power series must also be identical, since there is only one unique quotient resulting from a division process. Therefore no error remains.

In the case where  $H(s)$  is a transcendental function, we cannot hope to get identical series from the transcendental function and its rational function approximation, but we can show that the time-domain approximation must be close over certain intervals of time. We know from Laplace transformation theory that

$$\lim_{s \rightarrow 0} s H(s) = \lim_{t \rightarrow \infty} h(t) \quad (32)$$

where  $s H(s)$  is analytic on the axis of imaginaries and in the right half-plane (12).

For  $H(s) = b_0 + b_1 s + b_2 s^2 + b_3 s^3 + \dots$ ,  $s H(s)$  is analytic on the imaginary axis and right half-plane and, therefore, we may employ this theorem. As  $s \rightarrow 0$ , the terms involving higher powers of  $s$  become insignificant and, therefore, the behavior of  $h(t)$  for large  $t$  is fixed by the first few terms of the above power series. Since it is these terms whose coefficients we have matched exactly, this method assures us that our approximation in the time domain for large  $t$  is correct.

Next, we take care of the behavior of  $h(t)$  around  $t = 0$  by choosing the relative degree of the numerator and denominator polynomials of  $H^*(s) = P(s)/Q(s)$  to satisfy the requirements of the time function. Let us amplify this statement.

We may catalogue the various types of impulse responses according to their behavior as  $t \rightarrow 0$  (Fig. 15), which gives information about  $H(s)$  as  $s \rightarrow \infty$  through the initial value theorem

$$\lim_{t \rightarrow 0} h(t) = \lim_{s \rightarrow \infty} s H(s) \quad (33)$$

Since  $H^*(s)$  is a ratio of polynomials as  $s \rightarrow \infty$ , the whole function behaves as

$$\frac{p_n s^n}{q_m s^m}$$

where  $n$  and  $m$  are the degrees of the two polynomials; that is,  $s H^*(s)$  behaves as  $(p_n s^{n+1-m})/q_m$  for  $s \rightarrow \infty$ . We apply the constraint on  $H^*(s)$ , to take into account the behavior of  $h(t)$  for  $t \rightarrow 0$  by choosing the relative degree of the two polynomials  $P(s)$  and  $Q(s)$  in accordance with the catalogue of Fig. 15.

1. If  $h(t)$  starts out as a step, this means

$$\frac{P(s)}{Q(s)} \rightarrow \frac{1}{s} \quad \text{as } s \rightarrow \infty \quad (34)$$

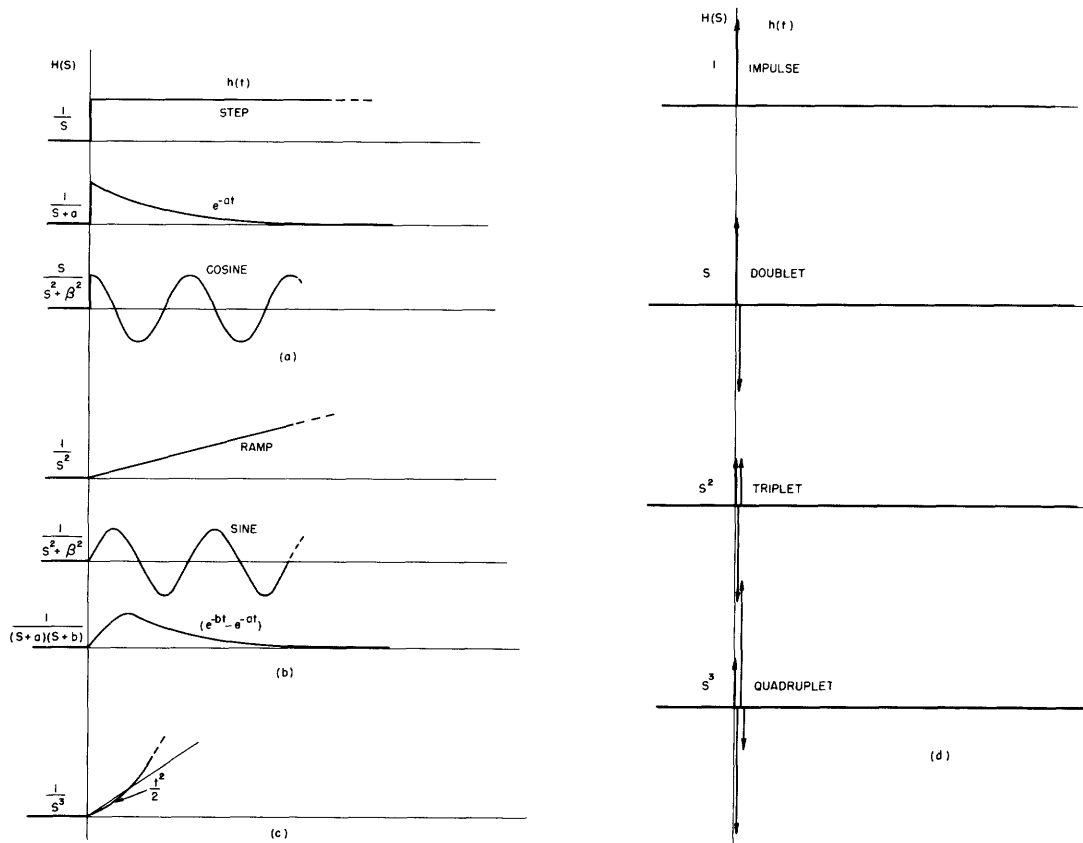


Fig. 15

and therefore  $n - m = -1$  or  $m = n + 1$ .

2. If  $h(t)$  begins with a linear slope like a ramp function,

$$\frac{P(s)}{Q(s)} \rightarrow \frac{1}{s^2} \quad \text{as } s \rightarrow \infty \quad (35)$$

and, therefore,  $n - m = 2$  or  $m = n + 2$ .

3. On the other hand, if  $h(t)$  contains impulses, doublets, or triplets, or the higher order singularity functions at  $t = 0$ , we may treat them separately from the proper fraction  $H^*(s)$  and write at once

$$H(s) = k_0 + k_1 s + k_2 s^2 + \dots + H^*(s) \quad (36)$$

where  $H^*(s) = P(s)/Q(s)$  (with the denominator polynomial of higher degree than the numerator polynomial and the relative degree which we denote by  $r = m - n$ ) is obtained as in case 1 and case 2 above. More specifically, we may obtain this relative degree  $r$  by a short calculation as follows:

a. Let the ordinate values of  $h(t)$ , near  $t = 0$ , be denoted by  $h_0, h_1, h_2, h_3, h_4, \dots$  with  $\Delta t$  small and where  $h_0$  refers to  $h(0)$ .

Then if  $h_0 = k_0$  where  $k_0 \neq 0$ ,

$$h_1 = k_1$$

$$h_2 = k_2$$

$$r = -1$$

because we recognize that  $h(t)$  starts out with a jump at  $t = 0$ .

b. If  $\frac{h_2}{h_1} = 2$

$$\frac{h_3}{h_1} = 3, \quad \text{and} \quad h_0 = 0,$$

$$r = -2$$

because this denotes a discontinuity in the slope at  $t = 0$ .

c. If  $\frac{h_2}{h_1} > 2$ , say  $2 + k$

$$\frac{h_3}{h_1} \leq 3 + 3k, \quad \text{and} \quad h_0 = 0,$$

$$r = -3$$

Let  $h(t) = k_1 t + k_2 t^2$  where  $k_1$  and  $k_2 = \text{constants}$ ,

$$h_1 = k_1 \Delta t + k_2 (\Delta t)^2$$

$$h_2 = k_1^2 \Delta t + k_2^4 (\Delta t)^2$$

$$h_3 = k_1^3 \Delta t + k_2^9 (\Delta t)^2$$

$$\frac{h_2}{h_1} = \frac{k_1^2 \Delta t + k_2^4 (\Delta t)^2}{k_1 \Delta t + k_2 (\Delta t)^2} = \frac{2 + k_2/k_1^4 \Delta t}{1 + k_2/k_1 \Delta t} = \frac{2 + 4 k_3}{1 + k_3} = 2 + k$$

$$\frac{h_3}{h_1} = \frac{k_1^3 \Delta t + k_2^9 (\Delta t)^2}{k_1 \Delta t + k_2 (\Delta t)^2} = \frac{3 + k_2/k_1^9 \Delta t}{1 + k_2/k_1 \Delta t} = \frac{3 + 9 k_3}{1 + k_3} = 3 + 3k$$

d. If  $h_0 = 0$ ,

$$\frac{h_2}{h_1} > 2, \quad \text{say} \quad \frac{h_2}{h_1} = 2 + k$$

$$\frac{h_3}{h_1} > 3 + 3k, \quad \text{say} \quad \frac{h_3}{h_1} = 3 + 3k + k_4$$

and

$$\frac{h_4}{h_1} \leq 4 + 6k + 4k_4; \quad r = -4$$

We may check by letting  $h(t) = at^3$

$$h_1 = a(\Delta t)^3$$

$$h_2 = a8(\Delta t)^3$$

$$h_3 = a27(\Delta t)^3$$

$$h_4 = a64(\Delta t)^3$$

$$\frac{h_2}{h_1} = 8 = 2 + 6, \quad k = 6$$

$$\frac{h_3}{h_1} = 27 = 3 + 18 + 6, \quad k_4 = 6 = 3 + 3k + k_4$$

$$\frac{h_4}{h_1} = 64 = 4 + 36 + 24 = 4 + 6k + 4k_4$$

We get these relations by assuming

$$h(t) = at^3 + bt^2 + ct \tag{37}$$

and going through the same process as in Part c, but with much more arithmetic.

The above examples are sufficient to illustrate the procedure for obtaining the relative degree  $r$  that assures us of the correct behavior for  $h^*(t)$ , the approximating function for  $t \rightarrow 0$ .

Thus we have fixed the behavior of  $h^*(t)$  for both small  $t$  and large  $t$  by Eqs. 32 and 33. Now we show how the function is controlled in the intervening interval.

We start from

$$H(s) = \int_0^\infty h(t) \exp(-st) dt \tag{38}$$

and use the power series for  $\exp(-st)$ ,  $(st)^2 < \infty$ , to get

$$H(s) = \int_0^\infty h(t) dt - s \int_0^\infty t h(t) dt + \frac{s^2}{2!} \int_0^\infty t^2 h(t) dt \dots \tag{39}$$

We see that by identifying this with Eq. 29

$$H(s) = b_0 + b_1 s + b_2 s^2 + b_3 s^3 + \dots$$

that  $b_0$  corresponds to the area under the  $h(t)$  curve,  $b_1$  corresponds to the first moment,  $b_2$  corresponds to the second moment, and so on. Thus we find that as we match  $b_0$ ,  $b_1$ ,  $b_2$ , and succeeding coefficients by Eq. 31, we are specifying more and more about the shape of the curve in the intermediate interval.

This also tells us that when we are trying to approximate a time function whose Laplace transform is a transcendental function by a  $H^*(s)$  in rational function form, we should get a better fit to the  $h(t)$  by using more of the coefficients  $b_0$ ,  $b_1$ ,  $b_2$ ,  $\dots$ ; in other words, by employing polynomials of higher degree satisfying the requirement on  $r$ , the relative degree. Hence we can reduce the error by using more complicated networks. We will now illustrate, by a few examples, how this method gives good approximations in the time domain for transcendental  $H(s)$  and precise results for rational function  $H(s)$ .

## 5.2 APPLICATION TO WELL-KNOWN CASES TO SHOW RAPIDITY AND SIMPLICITY OF COMPUTATION

(i)  $h(t) = \exp(-t)$

We know that this has the Laplace transform  $H(s) = 1/s+1$  for the system function. It is so simple that it could be guessed, but we include it to illustrate the computations for a short example where we do not need much arithmetic.

By a look at the shape of the curve we see that the relative degree of the numerator and denominator polynomials is  $n - m = -1$  and that there is just one pole. So we try

$$H^*(s) = \frac{p_0}{s + q_0} \quad (40)$$

since we may assume  $q_1 = 1$ , without loss of generality.

The power series is

$$H(s) = 1 - s + s^2 - s^3 + s^4 \dots$$

therefore

$$\frac{p_0}{s + q_0} = 1 - s + s^2 \dots \quad (41)$$

or

$$p_0 = q_0 + s(1 - q_0) + s^2(-1 + q_0) + \dots$$

$$p_0 = q_0$$

$$0 = 1 - q_0$$

or

$$q_0 = 1$$

$$p_0 = 1$$

therefore

$$H^*(s) = \frac{1}{s+1}$$

which is exactly equal to the true system function.

We will find that all equations that may be written from Eq. 41 are satisfied by the above  $p_0$  and  $q_0$ .

$$(ii) \quad h(t) = t \exp(-t)$$

The power series expansion is

$$H(s) = 1 - 2s + 3s^2 - 4s^3 + 5s^4 \dots \quad (42)$$

The relative degree  $r$  equals  $-2$  for this time function. Let

$$H^*(s) = \frac{a}{1 + cs + bs^2}$$

since one of the coefficients  $q_0$  may be set equal to 1 without loss of generality.

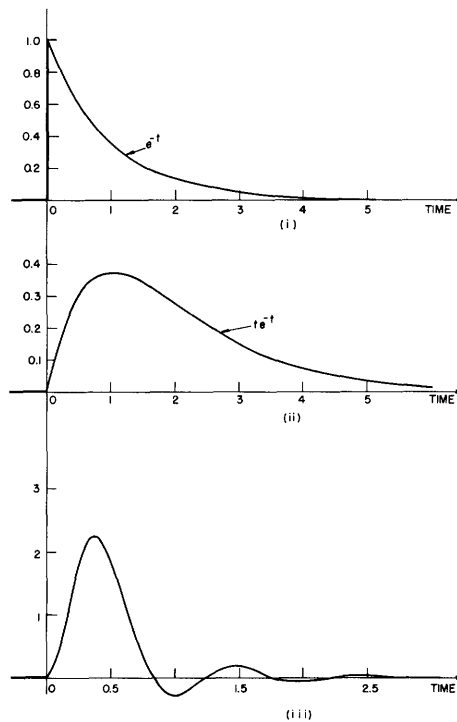


Fig. 16

$$\frac{a}{1 + cs + bs^2} = 1 - 2s + 3s^2 - 4s^3 + 5s^4 \dots \quad (43)$$

gives

$$a = 1$$

$$0 = c - 2$$

$$0 = b - 2c + 3$$

$$0 = -2b + 3c - 4$$

therefore

$$a = 1$$

$$c = 2$$

$$b = -3 + 4 = 1$$

This further satisfies  $-2b + 3c - 4 = 0$  and later equations. Therefore

$$H^*(s) = \frac{1}{s^2 + 2s + 1}$$

as may be checked from a table of Laplace transforms.

(iii) Impulse response as in Fig. 16(iii)

The power series expansion is

$$H(s) = 1 - 0.4333s + 0.12944s^2 - 0.03915s^3 + \dots$$

From

$$\frac{h_2}{h_1} > 2 = 2 + k$$

$$\frac{h_3}{h_1} < 3 + 3k$$

we obtain  $r = -3$ . Let

$$H^*(s) = \frac{a}{1 + bs + cs^2 + ds^3} \quad (44)$$

$$\frac{a}{1 + bs + cs^2 + ds^3} = 1 - 0.4333s + 0.12944s^2 - 0.03915s^3 + \dots \quad (45)$$

$$a = 1$$

$$0 = b - 0.4333$$

$$0 = c - 0.4333b + 0.12944$$

$$0 = d - 0.4333c + 0.12944b - 0.03915$$

Solving for the unknowns we get at once

$$a = 1$$

$$b = 0.4333$$

$$c = 0.4333b - 0.12944 = 0.05832$$

$$d = 0.00834$$

Therefore

$$H^*(s) = \frac{1}{1 + 0.4333s + 0.05832s^2 + 0.00834s^3}$$

which is identical to

$$H(s) = \frac{120}{s^3 + 7s^2 + 52s + 120} = \frac{120}{(s+3) [(s+2)^2 + 6^2]}$$

which is the system function of the network whose impulse response is shown in Fig. 16(iii).

$$h(t) = 120 [0.027 \exp(-3t) + 0.0274 \exp(-2t) \sin(6t - 1.41)]$$

as may be checked from the transform pair

$$\frac{k}{(s+\gamma) [(s+a)^2 + \beta^2]}$$

and its inverse Laplace transform

$$\frac{\exp(-\gamma t)}{(\gamma-a)^2 + \beta^2} + \frac{\exp(-at) \sin(\beta t - \psi)}{\beta \sqrt{[(\gamma-a)^2 + \beta^2]}}$$

with  $\psi = \tan^{-1} \left[ \frac{\beta}{\gamma-a} \right]$

These examples indicate that in the case of  $h(t)$  whose Laplace transforms are rational functions, we recover these precisely by using

$$H^*(s) = \frac{p_0 + p_1 s + \dots + p_n s^n}{q_0 + q_1 s + \dots + q_m s^m}$$

and the two time functions are identical.

We now go to a very important problem in time-domain synthesis: how to obtain an acceptable approximation to an impulse response whose Laplace transform is a transcendental function. By an acceptable approximation it is understood that the impulse response of the rational function network (which is a finite, lumped-parameter network, in contrast to the infinite or distributed-parameter network corresponding to the transcendental function) follows the specified impulse response within prescribed tolerances,



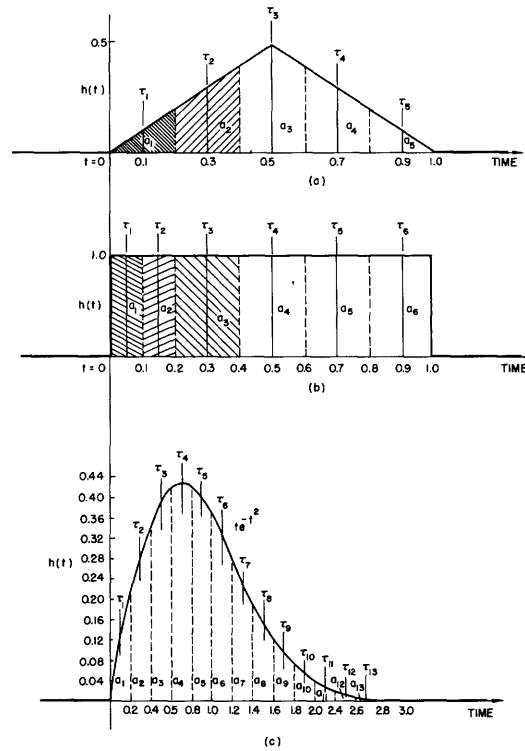


Fig. 17

and at the same time requires an economical number of elements in the network. We may say that the aim is to get as few poles in  $H^*(s) = P(s)/Q(s)$  to fit the specified impulse response with the prescribed tolerances. We fulfill this in section 5.3; but now we wish to attend to the question of the accuracy of the coefficients in the power series expansion of a transcendental function when we calculate them by means of the Dirichlet series, using the areas under the time-function curve.

a. Let us consider the case of a triangular impulse response with unit slope and discontinuities in the slope at  $t = 0, 0.5$ , and  $1$ . We know that the transcendental Laplace transform for it is

$$H(s) = \frac{1}{s^2} [1 - \exp(-0.5s)]^2$$

whose power series expansion is

$$\begin{aligned} H(s) = & 0.25 - 0.125s + 0.0364583s^2 - 0.0078125s^3 + 0.00134548s^4 \\ & - 0.000195313s^5 + 0.000024607824s^6 \dots \end{aligned} \quad (46)$$

Now we calculate the power series expansion from the Dirichlet series. Dividing time

into the intervals shown in Fig. 17(a) we have the areas

$\tau_n$	$a_n$
0.1	0.02
0.3	0.06
0.5	0.09
0.7	0.06
0.9	0.02
	0

for all  $t > 1$ .

$$H(s) = \sum_{n=1}^5 a_n \exp(-\tau_n s)$$

$$\begin{aligned}
H(s) = & 0.02 \left[ 1 - (0.1s) + \frac{(0.1s)^2}{2!} - \frac{(0.1s)^3}{3!} + \frac{(0.1s)^4}{4!} \dots \right] \\
& + 0.06 \left[ 1 - (0.3s) + \frac{(0.3s)^2}{2!} - \frac{(0.3s)^3}{3!} + \frac{(0.3s)^4}{4!} \dots \right] \\
& + 0.09 \left[ 1 - (0.5s) + \frac{(0.5s)^2}{2!} - \frac{(0.5s)^3}{3!} + \frac{(0.5s)^4}{4!} \dots \right] \\
& + 0.06 \left[ 1 - (0.7s) + \frac{(0.7s)^2}{2!} - \frac{(0.7s)^3}{3!} + \frac{(0.7s)^4}{4!} \dots \right] \\
& + 0.02 \left[ 1 - (0.9s) + \frac{(0.9s)^2}{2!} - \frac{(0.9s)^3}{3!} + \frac{(0.9s)^4}{4!} \dots \right] \quad (47)
\end{aligned}$$

$$H(s) = 0.25 - 0.125s + 0.03685s^2 - 0.008008s^3 + 0.00013835s^4 \dots \quad (48)$$

Comparison of the two series in Eqs. 46 and 48 shows that even with the intervals chosen at 20 percent of the total duration, the accuracy is high; the error being only 1 percent for the first, second, and third coefficients, 2.5 percent for the fourth, and 2.8 percent for the fifth coefficient. With smaller intervals the accuracy increases and becomes perfect as  $n \rightarrow \infty$ .

b. We may choose the intervals of different size and still maintain fair accuracy. For example, for a rectangular pulse divided as in Fig. 17(b) we have the following results

$\tau_n$	$a_n$
0.05	0.1
0.15	0.1
0.3	0.2
0.5	0.2
0.7	0.2
0.9	0.2

$$\begin{aligned}
H(s) &= \sum_{n=1}^6 a_n \exp(-\tau_n s) \\
&= 0.1 \left[ 1 - (0.05s) + \frac{(0.05s)^2}{2!} - \frac{(0.05s)^3}{3!} + \frac{(0.05s)^4}{4!} \dots \right] \\
&\quad + 0.1 \left[ 1 - (0.15s) + \frac{(0.15s)^2}{2!} - \frac{(0.15s)^3}{3!} + \frac{(0.15s)^4}{4!} \dots \right] \\
&\quad + 0.2 \left[ 1 - (0.3s) + \frac{(0.3s)^2}{2!} - \frac{(0.3s)^3}{3!} + \frac{(0.3s)^4}{4!} \dots \right] \\
&\quad + 0.2 \left[ 1 - (0.5s) + \frac{(0.5s)^2}{2!} - \frac{(0.5s)^3}{3!} + \frac{(0.5s)^4}{4!} \dots \right] \\
&\quad + 0.2 \left[ 1 - (0.7s) + \frac{(0.7s)^2}{2!} - \frac{(0.7s)^3}{3!} + \frac{(0.7s)^4}{4!} \dots \right] \\
&\quad + 0.2 \left[ 1 - (0.9s) + \frac{(0.9s)^2}{2!} - \frac{(0.9s)^3}{3!} + \frac{(0.9s)^4}{4!} \dots \right]
\end{aligned}$$

$$H(s) = 1 - 0.5s + 0.16524s^2 - 0.04086s^3 + 0.008049s^4 \dots \quad (49)$$

The Laplace transform of the rectangular impulse response is

$$H(s) = \frac{1}{s} [1 - \exp(-s)] \quad (50)$$

whose power series expansion is

$$H(s) = 1 - 0.5s + 0.16666s^2 - 0.0416666s^3 + 0.0083333s^4 \dots \quad (51)$$

From Eqs. 49 and 51, the accuracy is seen to be good, even for unequal and wide spacing. The errors are less than 1 percent for the first, second, and third coefficients, 2 percent

for the fourth and 4 percent for the fifth.

c. We now take a different type of transcendental function, one whose time function is  $t \exp(-t^2)$ . The transcendental function is

$$H(s) = \mathcal{L}[t \exp(-t^2)] = -\frac{d}{ds} \left[ \frac{(\pi)^{1/2}}{2} \exp(s^2/4) \operatorname{cerf} \frac{s}{2} \right] \quad (52)$$

where

$$\begin{aligned} \operatorname{cerf} \frac{s}{2} &= \frac{2}{(\pi)^{1/2}} \int_{s/2}^{\infty} \exp(-x^2) dx \\ &= \frac{2}{(\pi)^{1/2}} \left( 1 - \operatorname{erf} \frac{s}{2} \right) \end{aligned} \quad (53)$$

$$\operatorname{erf} \frac{s}{2} = \frac{2}{(\pi)^{1/2}} \left( \frac{s}{2} - \frac{s^3}{2^3 \times 3} + \frac{s^5}{2^5 \times 5 \times 2!} - \frac{s^7}{2^7 \times 7 \times 3!} + \dots \right) \quad (54)$$

$$\begin{aligned} H(s) &= \exp(s^2/4) \left[ 1 - \frac{(\pi)^{1/2}}{2} s + \frac{s^2}{4} - \frac{s^4}{96} + \frac{s^6}{1920} - \dots \right] \\ &= 0.5 - 0.443115s + 0.25s^2 - 0.11077875s^3 + 0.041666s^4 \\ &\quad - 0.01384734s^5 + 0.0041666s^6 - 0.0011539s^7 + \dots \end{aligned} \quad (55)$$

For obtaining the Dirichlet series representation we have

$\tau_n$	$a_n$
0.1	0.0192
0.3	0.0537
0.5	0.0767
0.7	0.0856
0.9	0.0796
1.1	0.0655
1.3	0.0481
1.5	0.0323
1.7	0.0194
1.9	0.0107
2.1	0.0053
2.3	0.0024
2.5	0.0007

The power series expansion is then obtained as

$$H(s) = 0.4985 - 0.44123s + 0.24604s^2 - 0.1086s^3 + 0.03995s^4 - 0.01308s^5 + \dots \quad (56)$$

The accuracy in this case is also good. The error in the first two coefficients is less than 1 percent; and 2 percent in the third and fourth coefficients.

### 5.3 FINDING THE RATIONAL FUNCTION APPROXIMATION TO A TRANSCENDENTAL FUNCTION BY USING THE TIME FUNCTION $h(t)$

In many of the practical applications of network synthesis the problem is to find a finite, linear, passive network that will produce an impulse response whose shape is such that its Laplace transform is a transcendental function. In terms of the theory of functions of a complex variable, a transcendental function may have an essential singularity at infinity or an infinite number of poles in the finite  $s$ -plane. This means that to get them exactly in network form entails an infinite number of elements or distributed parameters. Therefore it is of the utmost practical importance to obtain a rational function approximation to the transcendental system function. This would enable us to realize the impulse response to within the prescribed tolerances by means of a finite, lumped-parameter network.

One of the difficult impulse responses to approximate in the time domain is the triangular pulse shown in Fig. 17(a).

Its Laplace transform involves two transcendentals  $\exp(-0.5s)$  and  $\exp(-s)$  because

$$H(s) = \frac{1}{s^2} [1 - 2 \exp(-0.5s) + \exp(-s)] \quad (57)$$

If we use the power series expansion method it is quite a simple matter to get a rational function approximation. From the slope as  $t \rightarrow 0$ , we get the relative degree of numerator and denominator polynomials  $r = n - m = -2$ .

For a first example let

$$H_1(s) = \frac{a + bs}{1 + cs + ds^2 + es^3} \quad (58)$$

$$H(s) = 0.25 - 0.125s + 0.0364583s^2 - 0.0078125s^3 + 0.001345486s^4 \dots \quad (46)$$

From Eqs. 58 and 46 we have the set

$$\begin{aligned} a &= 0.25 \\ b &= -0.125 + 0.25c \\ 0 &= 0.0364583 - 0.125c + 0.25d \\ 0 &= -0.0078125 + 0.0364583c - 0.125d + 0.25e \\ 0 &= 0.001345486 - 0.0078125c + 0.0364583d - 0.125e \end{aligned}$$

We rewrite the last two equations as

$$\begin{array}{rcl}
 -e + 0.291666d - 0.0625c & = & -0.01076389 \\
 e - 0.5d & + & 0.145833c = 0.03125 \\
 \hline
 -0.208334d + 0.08333c & = & 0.02048611
 \end{array}$$

Combining this with

$$0.25d - 0.125c = 0.0364583$$

we have

$$-0.0208337c = -0.0098959$$

$$c = 0.474995$$

Therefore

$$d = 0.09166422$$

Therefore

$$e = 0.03125 - 0.1458332 \times 0.474995 + 0.5 \times 0.0916642 = 0.007812$$

$$b = -0.125 + 0.25 \times 0.474995 = -0.00625$$

$$a = 0.25$$

Therefore

$$H_1(s) = \frac{0.25 - 0.00625s}{1 + 0.47995s + 0.09166422s^2 + 0.007812s^3} \quad (60)$$

or

$$H_1(s) = \frac{-0.8000512 (s - 40)}{(s + 4.5555) [(s + 3.588966)^2 + 3.900901^2]}$$

which has one negative real pole and a pair of complex poles. This rational function has

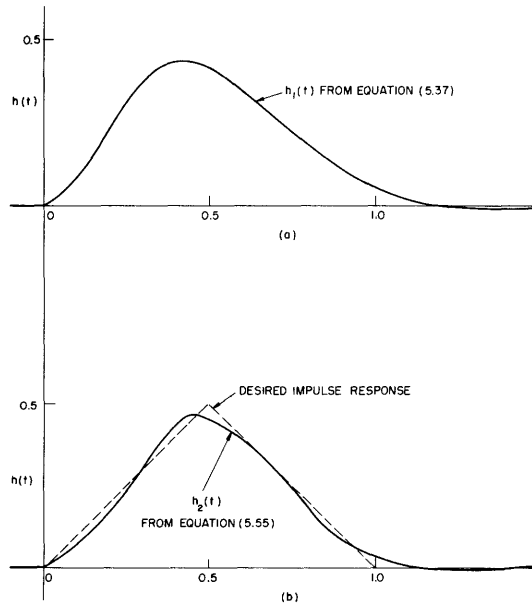


Fig. 18

the inverse Laplace transform

$$h_1(t) = 2.20705859 \exp(-4.555t) + 2.2333597 \exp(-3.588966t) \sin(3.900901t - 81.2^\circ) \quad (61)$$

This has been plotted in Fig. 18(a). We see that the approximating curve does decay rapidly as  $t \rightarrow \infty$  and has the same area as predicted in section 5.1.

To get a better approximation, let us now use a fourth-degree denominator. Because  $r = -2$ , the numerator will be of second degree. So, let

$$H_2(s) = \frac{a + bs + cs^2}{1 + ds + es^2 + fs^3 + gs^4} \quad (62)$$

From Eqs. 46 and 62 we get the equations

$$a = 0.25 \quad (63)$$

$$b = -0.125 + 0.25d \quad (64)$$

$$c = 0.03645836 - 0.125d + 0.25e \quad (65)$$

$$0 = -0.0078125 + 0.03645836d - 0.125e + 0.25f \quad (66)$$

$$0 = 0.001345486 - 0.0078125d + 0.03645836e - 0.125f + 0.25g \quad (67)$$

$$0 = -0.000195313 + 0.001345486d - 0.0078125e + 0.03645836f - 0.125g \quad (68)$$

$$0 = 0.000024607824 - 0.000195313d + 0.001345486e - 0.0078125f + 0.0364584g \quad (69)$$

From Eqs. 69 and 68

$$0.077381f - 0.025595e + 0.00540674d = 0.00088755 \quad (70)$$

From Eqs. 68 and 67

$$-0.208333f + 0.08333e - 0.020486d = -0.00381944 \quad (71)$$

From Eqs. 70, 71, and 62

$$f - 0.330765e + 0.069871674d = 0.01146987 \quad (72)$$

$$-f + 0.4e - 0.098332957d = -0.0183333 \quad (73)$$

$$f - 0.5e + 0.14583344d = 0.03125 \quad (74)$$

we get

$$-0.1e + 0.0475005d = 0.01291666 \quad (75)$$

$$0.069235e - 0.0284613d = -0.00686347 \quad (76)$$

$$0.00442567d = 0.00207938$$

or

$$d = 0.46984524$$

therefore

$$e = 0.094012178$$

$$f = 0.03125 - 0.14583344 \times 0.46984524 + 0.5 \times 0.094012178 = 0.009736942$$

$$g = -0.005381944 + 0.03125 \times 0.46984524 - 0.14583344 \times 0.094012178 \\ + 0.5 \times 0.009736942 = 0.00045907443$$

$$c = 0.03645836 - 0.125 \times 0.46984524 + 0.25 \times 0.094012178 = 0.0012307495$$

$$b = -0.125 + 0.25 \times 0.46984524 = -0.00753869$$

$$H_2(s) = \frac{0.25 - 0.00753869s + 0.0012307495s^2}{1 + 0.46984524s + 0.094012178s^2 + 0.009736942s^3 + 0.000459074s^4} \quad (77)$$

$$= \frac{0.25 - 0.00753869s + 0.0012307495s^2}{0.00045907443 (s^2 + 12.1107215s + 40.266915)(s^2 + 9.09914s + 54.0964213)}$$

$$= 2178.296 \left[ \frac{0.0002764354s + 0.012069782695}{(s + 6.0553607)^2 + 1.89725^2} - \frac{0.0002764354s + 0.01000653}{(s + 4.54957)^2 + 5.779085^2} \right] \quad (78)$$

This rational function has 2 pairs of complex poles. Its inverse Laplace transform is

$$h_2(t) = 11.951026973 \exp(-6.05536t) \sin(1.89725t + 2.9^\circ) \\ - 3.352218166 \exp(-4.54957t) \sin(5.779085t + 10.3^\circ) \quad (79)$$

For just two pairs of poles the time-domain approximation can be considered fairly good (well within 10 percent). There is very little oscillation, as may be seen from Fig. 18(b).

The fact that, by using more coefficients in the power series expansion Eq. 46, we get a time function whose shape is closer to the triangle, confirms the statement made in section 5.1 on theoretical grounds.

We shall see the same convergence towards the specified impulse response as we use more poles in the approximation, in the next example of a transcendental system function; namely, the one of section 5.2(c).

$$H(s) = \frac{-d}{ds} \left[ \frac{(\pi)^{1/2}}{2} \exp(s^2/4) \operatorname{cerf} \frac{s}{2} \right] \quad (52)$$

This differs from the previous case in that its transcendental components are not the result of simple delay factors like  $\exp(-s)$ ; and so it does not lend itself to the  $s$ -domain approximations known as the Padé functions (40). However, by our method (using the power series expansion) a good approximation is easily obtained, with great economy of poles.

Let us first try to approximate it with just one pair of complex poles.

$$H_1(s) = \frac{a}{1 + bs + cs^2} \quad (80)$$



since the slope around  $t \rightarrow 0$  is linear.

From Eq. 55

$$H(s) = 0.5 - 0.443115s + 0.25s^2 \dots \quad (81)$$

Therefore, by Eqs. 80 and 81

$$a = 0.5 \quad (82)$$

$$0 = 0.443115 + 0.5b \quad (83)$$

$$0 = 0.25 - 0.443115b + 0.5c \quad (84)$$

Therefore

$$b = 0.88623$$

$$c = 0.2854036$$

$$H_1(s) = \frac{0.5}{1 + 0.88623s + 0.2854036s^2} \quad (85)$$

Therefore

$$h_1(t) = 0.5 \exp(-1.552591t) \sin 1.045712t \quad (86)$$

This is plotted in Fig. 19(a) and shows the same shape as the required impulse response. We get a very close fit by trying next a rational function approximation with one real negative pole and a pair of complex poles, that is, we choose our denominator polynomial of third degree. Let

$$H_2(s) = \frac{a + bs}{1 + cs + ds^2 + es^3} \quad (87)$$

From Eqs. 87 and 55 we get the set of equations

$$a = 0.5 \quad (88)$$

$$b = -0.443115 + 0.5c \quad (89)$$

$$0 = 0.25 - 0.443115c + 0.5d \quad (90)$$

$$0 = -0.11077875 + 0.25c - 0.443115d + 0.5e \quad (91)$$

$$0 = 0.041666 - 0.11077875c + 0.25d - 0.443115e \quad (92)$$

From Eqs. 91 and 92

$$e - 0.886230d + 0.5c = 0.2215575 \quad (93)$$

$$-e + 0.564187625d - 0.25c = -0.09403127 \quad (94)$$

$$\begin{array}{rcl} -0.3220424d & + & 0.25c = 0.127526 \end{array} \quad (95)$$

$$\begin{array}{rcl} +0.3220424d & - & 0.2854036c = -0.1610212 \end{array} \quad (90)$$

$$\begin{array}{rcl} & & -0.0354036c = -0.0334952 \end{array}$$

Therefore

$$c = 0.946095877$$

$$d = 0.338458445$$

$$e = 0.22155750 - 0.5 \times 0.94609588 + 0.88623 \times 0.338458445 = 0.048461588$$

$$b = 0.5 \times 0.94609588 - 0.443115 = 0.02993294$$

$$a = 0.5$$

Therefore

$$\begin{aligned} H_2(s) &= \frac{0.5 + 0.02993294s}{1 + 0.946095877s + 0.338458445s^2 + 0.048461588s^3} \\ &= \frac{0.5 + 0.02993294s}{(1 + 0.414968s)(1 + 0.531128s + 0.115384s^2)} \\ &= 0.6251517 \left[ \frac{4.2274374}{s + 2.40982437} - \frac{4.2274374 (s + 1.9567197)}{(s + 2.301547)^2 + 1.83563^2} \right] \end{aligned} \quad (92)$$

From the inverse Laplace transform

$$\begin{aligned} h_2(t) &= 2.64278968 \exp(-2.40982437t) \\ &\quad - 2.689033424 \exp(-2.301547t) \sin(1.83563t - 79.36^\circ) \end{aligned} \quad (93)$$

We plot the response of this network in Fig. 19(b). Comparison with the superimposed  $t \cdot \exp(-t^2)$  plot shows that we have achieved a close approximation to within 5 percent of the desired impulse response, proving this to be quite an economical method.

#### 5.4 CALCULATING $H(s)$ FROM $h(t)$ ALGEBRAICALLY BY THE TIME DOMAIN OPERATOR REPRESENTATION FOR $s$ , $s^2$ , $s^n$

We saw in section 3.3 how the doublet in the time domain, represented in the time sequence form as  $\{1/\Delta t, -1/\Delta t, 0, 0, 0, \dots\}$ , performed differentiation on another time sequence. It was the analog of

$$F_i(s) \cdot H(s), \quad \text{where } H(s) = s \quad (94)$$

This gave rise to the idea of representing more complicated  $H(s)$ , where

$$H(s) = \frac{P(s)}{Q(s)} = \frac{p_0 + p_1s + \dots + p_ns^n}{q_0 + q_1s + \dots + q_ms^m} \quad (95)$$

by similar time sequences and determining the unknown coefficients directly from the time sequence of the impulse response. In practice, this works well for finding  $H(s)$  from  $h(t)$  in those cases where the system function turns out to be a rational function, but is not so powerful as the power series expansion method when dealing with  $h(t)$  whose Laplace transforms are transcendental functions.

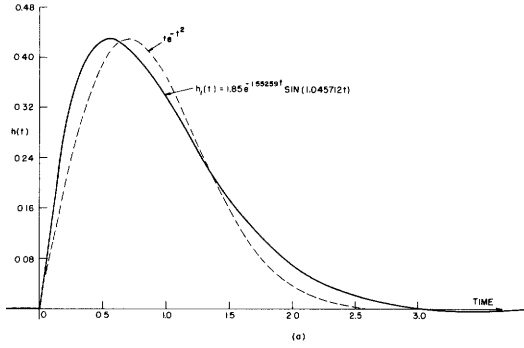


Fig. 19(a)

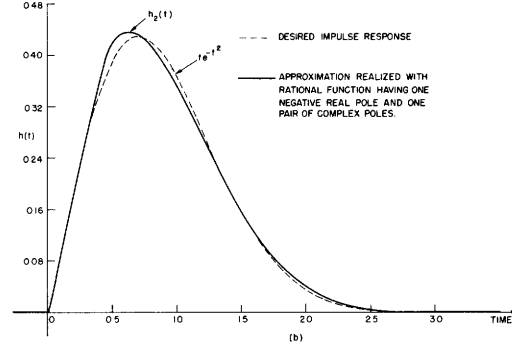


Fig. 19(b)

However, this new approach seems much more interesting in that it indicates something of a process by which a system function operates on the input to yield the output. For the sake of this insight we present this method. Moreover, by using the rational function approximation to  $\exp(-s)$ , it is possible to get around the limitation of the method when dealing with delayed time responses.

a. For the present, we shall describe the method for impulse responses whose Laplace transforms are rational functions. For this purpose let us use the simple impulse response

$$h(t) = 0.4 \exp(-2.5t)$$

to bring out the principle clearly.

This has a jump at  $t = 0$ , therefore the rational function will have a denominator of one degree higher than the numerator.

$$H_1(s) = \frac{x}{s + y} \quad (96)$$

Since  $h(t)$  is an impulse response, it means that if we use an impulse as an input, the output of the network will be  $h(t)$ . Thus we may write in time sequence form, remembering the theoretical justification of this representation as given in section 3.2,

$$\{h_1, h_2, h_3, \dots\} = \{1/\Delta t, 0, 0, 0, \dots\} \{H(s)\} \quad (97)$$

where  $\{1/\Delta t, 0, 0, \dots\}$  represents the impulse input and  $\{H(s)\}$  is written as

$$\frac{\{x, 0, 0, 0, \dots\}}{\{y + \frac{1}{\Delta t}, -\frac{1}{\Delta t}, 0, 0, \dots\}}$$

That is, we write  $\{1/\Delta t, -1/\Delta t, 0, 0, 0, \dots\}$  for  $s$ , by intuition (since  $s$  represents differentiation in the complex frequency plane, and the doublet represents the same in the time domain we may interchange these representations, depending on whether we are

operating on time functions or on s functions).

Here, since we want to perform all of our operations in the time domain, we employ the doublet. Using  $\Delta t = 0.02$ , we have

$$\{0.38049, 0.36194, 0.34428, \dots\} = \left\{ \frac{1}{0.02}, 0, 0, \dots \right\} \frac{\{x, 0, 0, \dots\}}{\left\{ y + \frac{1}{0.02}, -\frac{1}{0.02}, 0, 0, \dots \right\}}$$

or

$$\left\{ \frac{0.38049}{0.02} + 0.38049y, \frac{0.36194}{0.02} + 0.36194y - \frac{0.38049}{0.02}, \dots \right\} = \left\{ \frac{x}{0.02}, 0, 0, 0, \dots \right\} \quad (98)$$

Equating coefficients of corresponding terms, we obtain

$$\frac{0.38049}{0.02} + 0.38049y = \frac{x}{0.02} \quad (99)$$

$$\frac{0.36194}{0.02} + 0.36194y - \frac{0.38049}{0.02} = 0 \quad (100)$$

Therefore

$$y = \frac{0.38049 - 0.36194}{0.02 \times 0.36194} = 2.56$$

$$x = 0.38049 + 0.38049 \times 0.02 \times 2.56 = 0.3995$$

Therefore

$$H_1(s) = \frac{0.3995}{s + 2.56}$$

which is correct within 2 percent. If we have more complicated time functions which require polynomials  $P(s)$  and  $Q(s)$  of higher degree in  $s$ , we use the following time sequences for

$$s^2 \rightarrow \left\{ \frac{1}{\Delta t^2}, \frac{-2}{\Delta t^2}, \frac{1}{\Delta t^2}, 0, 0, 0, \dots \right\}$$

$$s^3 \rightarrow \left\{ \frac{1}{\Delta t^3}, \frac{-3}{\Delta t^3}, \frac{3}{\Delta t^3}, \frac{-1}{\Delta t^3}, 0, 0, 0, \dots \right\}$$

$$s^4 \rightarrow \left\{ \frac{1}{\Delta t^4}, \frac{-4}{\Delta t^4}, \frac{6}{\Delta t^4}, \frac{-4}{\Delta t^4}, \frac{1}{\Delta t^4}, 0, 0, \dots \right\}$$

and so on. For example, if

$$P(s) = p_0 + p_1s + p_2s^2 + p_3s^3$$

we write it as a time-domain operator in the form

$$\left\{ \left( p_0 + \frac{p_1}{\Delta t} + \frac{p_2}{\Delta t^2} + \frac{p_3}{\Delta t^3} \right), \left( \frac{-p_1}{\Delta t} - \frac{2p_2}{\Delta t^2} - \frac{3p_3}{\Delta t^3} \right), \left( \frac{p_2}{\Delta t^2} + \frac{3p_3}{\Delta t^3} \right), -\frac{p_3}{\Delta t^3}, 0, 0 \right\}$$

We may note that when writing  $H(s)$ , the system function in time sequence form, the individual terms are the areas of the impulse, doublet, triplet, and so on. However, the terms of the impulse response sequence  $\{h_1, h_2, h_3, \dots\}$  and of the impulse input sequence  $\{1/\Delta t, 0, 0, 0, \dots\}$  are ordinate values using some value of  $\Delta t \rightarrow 0$ . This is necessary because convolution (which is what  $\{f\} \{h\}_A$  is approximating) involves an integration, as pointed out in section 3.1.

Before going on to a way of obtaining rational function approximations for transcendental system functions by this method, we include a short note on the peculiar phenomenon of a certain class of ordinary-looking input time functions going into an unstable network and emerging without displaying any positive exponential characteristics. Let us first explain this in the language of Laplace transforms.

An unstable network has a pole in the right-half plane. Let its system function be

$$H(s) = \frac{1}{(s-a)} R(s) \quad a > 0 \quad (101)$$

where  $R(s)$  is the "well-behaved" part, that is, all of its poles are in the left-half plane, and for our present purpose we need not worry about them. Then, in general,

$$F_o(s) = F_i(s) \cdot H(s) \quad (102)$$

$$= F_i(s) \frac{1}{(s-a)} R(s) \quad (103)$$

and, therefore,  $F_o(s)$  will contain a pole in the right-half plane and be unstable except in one unusual case: if  $F_i(s)$  has a zero in the same location in the right-half plane. Then that zero of  $F_i(s)$  and that pole of  $H(s)$  will cancel, and  $F_o(s)$  will no longer contain that pole. Thus  $f_o(t)$  will have no positive exponentials.

Since in the time domain (as we illustrated in section 4.2) the zeros of the transfer function do not show up very prominently, it has been a matter of curiosity to see why this exceptional phenomenon takes place in the time domain. Using the doublet representation of  $s$ , we may offer a picture of the process.

In Eq. 103 we notice that it does not make any difference to the output which function we regard as the input and which we regard as the system function. Since this is also true in the convolution integral, we may regard the term containing the right-hand zero of  $F_i(s)$  as operating on the positive exponential term in  $h(t)$ . That is, we consider what happens when  $\exp(at)$  is operated upon by something whose Laplace transform contains a zero at  $s = a$ , thus

$$\{\exp(at)\} \{(s-a)\}$$

where  $\{\exp(at)\}$  denotes the time sequence corresponding to  $\exp(at)$  and  $\{s-a\}$  denotes the time-sequence operator consisting of a differentiator and an amplifier of amplification  $-a$ . Since the derivative of  $\exp(at)$  is  $a \exp(at)$ , and the amplified term is  $-a \exp(at)$ , the net output due to the positive exponential component  $\exp(at)$  and  $(s-a)$  is zero. Therefore, the total output response will not contain positive exponential terms even though the network is unstable.

b. We now take up the problem of finding a rational function approximation for network system functions that involve transcendentals due to delay factors  $\exp(-s)$ . We shall consider the triangular impulse response that we dealt with previously by the power series expansion method in section 5.3. Its Laplace transform is

$$\frac{1}{s^2} [1 - 2\exp(-s) + \exp(-2s)] \quad (104)$$

for discontinuities in the slope at  $t = 1$  and  $t = 2$ .

This we can also obtain by an inspection of the curve. At  $t = 0$ , the response rises linearly, and if we use the doublet and triplet representations for  $s$  and  $s^2$  and write, analogous to Eq. 96,

$$h_1(s) = \frac{x}{s^2 + ys + z} \quad (105)$$

and carry out the procedure with  $t = 0.1$ ,

$$\begin{aligned} \{0.1, 0.2, 0.3, \dots\} &= \left\{ \frac{1}{0.1}, 0, 0, 0, \dots \right\} \\ &\times \frac{\{x, 0, 0, 0, 0, \dots\}}{\left\{ z + \frac{y}{0.1} + \frac{1}{(0.1)^2}, \frac{-y}{0.1} - \frac{2}{(0.1)^2}, \frac{1}{(0.1)^2}, 0, 0, \dots \right\}} \end{aligned}$$

or

$$\{0.1, 0.2, 0.3, \dots\} \{z + 10y + 100, -10y - 200, 100, 0, 0, \dots\} = \{10x, 0, 0, 0, \dots\}$$

Therefore

$$0.1z + y + 10 = 10x$$

$$0.2z + 2y + 20 - y - 20 = 0$$

$$0.3z + 3y + 30 - 2y - 40 + 10 = 0$$

Therefore

$$z = 0$$

$$y = 0$$

$$x = 1$$

Therefore  $H_1(s) = 1/s^2$ , which is quite apparent in this case, but may not always be

obtained directly in other cases. This method will always give the  $H_1(s)$  for the start of the response.

This representation of  $s$  and  $s^2$  is good only as  $\Delta t \rightarrow 0$ . Therefore when we get out to  $t = 1$ ,  $H_1(s)$  will give a curve that keeps rising with linear unit slope. But now the response suddenly changes its slope; and so we recognize that a delayed response has now started. Since its slope must be  $-2$  in order for the net slope to be  $-1$ , the system function for this delayed response must be

$$H_2(s) = \exp(-s) \left( \frac{-2}{s^2} \right) \quad (106)$$

These two system functions  $H_1(s)$  and  $H_2(s)$  keep the response going down with unit negative slope until  $t = 1$ , when a second delayed response that levels out the total response starts acting. Hence, the system function for this second delayed response is

$$H_3(s) = \exp(-2s) \left( \frac{1}{s^2} \right) \quad (107)$$

To get  $H(s)$ , the complete system function for the whole triangular response, in rational function form, we need a rational function approximation for  $\exp(-s)$ . Kautz (27) has done intensive investigation in this field and has come up with the Padé functions. The best approximation in the time domain, for a reasonable number of poles in the system function, appears to be the group

$$P_{25}(s), P_{36}(s), P_{47}(s), P_{58}(s)$$

where the first subscript refers to the degree of the numerator polynomial and the second to the degree of the denominator polynomial. We shall use  $P_{36}(s)$  to illustrate the method. This will give a network system function with six pairs of complex poles and a time-domain approximation within 10-percent tolerances.

The general formula for a Padé function is given as (27)

$$P_{\mu\nu} = \frac{1 - \frac{\mu}{\mu+\nu} \frac{s}{1!} + \frac{\mu(\mu-1)}{(\mu+\nu)(\mu+\nu-1)} \frac{s^2}{2!} + \dots + \frac{(-1)^\mu \mu(\mu-1) \dots 2 \cdot 1}{(\mu+\nu)(\mu+\nu-1)(\nu+1)} \frac{s^\mu}{\mu!}}{1 + \frac{\nu}{\mu+\nu} \frac{s}{1!} + \frac{\nu(\nu-1)}{(\mu+\nu)(\mu+\nu-1)} \frac{s^2}{2!} + \dots + \frac{\nu(\nu-1) \dots 2 \cdot 1}{(\mu+\nu)(\mu+\nu-1) \dots (\mu+1)} \frac{s^\nu}{\nu!}} \quad (108)$$

Therefore for  $\exp(-s)$ , using  $P_{36}(s)$ , we have

$$\exp(-s) \approx \frac{120 (-s^3 + 21s^2 - 168s + 504)}{s^6 + 24s^5 + 300s^4 + 2400s^3 + 12600s^2 + 40320s + 60480} \quad (109)$$

and for  $\exp(-2s)$

$$\exp(-2s) \approx \frac{120(-8s^3 + 84s^2 - 336s + 504)}{64s^6 + 768s^5 + 4800s^4 + 11200s^3 + 50400s^2 + 80640s + 60480} \quad (110)$$

Thus our system function, Eq. 104, becomes

$$\begin{aligned} & (64s^{10} + 2304s^9 + 42,432s^8 + 524,800s^7 + 4,257,600s^6 \\ & + 28,233,600s^5 + 141,214,080s^4 + 587,750,400s^3 \\ & H^*(s) = \frac{+1,239,974,400s^2 + 4,165,862,400s + 3,696,537,600}{(s^6 + 24s^5 + 300s^4 + 2400s^3 + 12600s^2 + 40320s + 60480)} \quad (111) \\ & \times (64s^6 + 768s^5 + 4800s^4 + 11200s^3 + 50400s^2 + 80640s + 60480) \end{aligned}$$

This rational function with six pairs of complex poles gives an impulse response that is shown in Fig. 20. Comparison with the example in section 5.3 for a similar triangular response shows that the power series expansion method achieves an approximation in the time domain within the same tolerances with much greater economy of poles (which means greater simplicity in the network). Moreover, since the Padé functions are

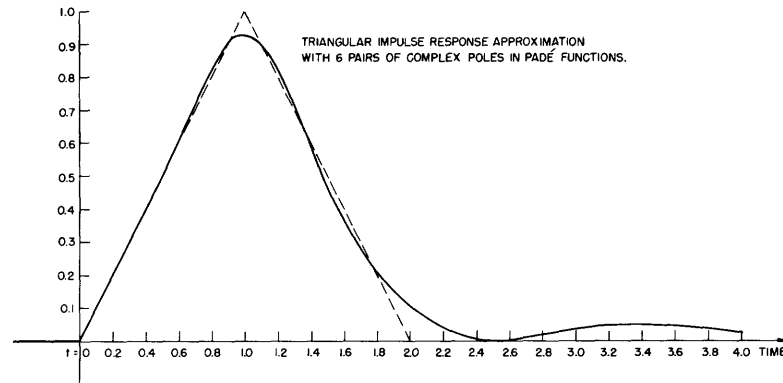


Fig. 20

essentially only approximations in the  $s$  domain, there is no guarantee of a good approximation in the time domain. This is borne out by the fact that the Padé functions  $P_{26}(s)$  and  $P_{56}(s)$ , which have the same number of poles as  $P_{36}(s)$ , have much poorer approximations in the time domain than the time function corresponding to  $P_{36}(s)$ .

#### 5.5 OBTAINING $H(s)$ FROM $h(t)$ BY USING THE TIME DOMAIN OPERATOR REPRESENTATION FOR $1/s$ , $1/s^2$ , $1/s^n$

Finally we present a method which is similar to the previous ones in that it involves only the solution of a set of linear, simultaneous, algebraic equations; and in that the concept is of a time domain representation; but in this case we use the integrating



operators  $1/s$ ,  $1/s^2$ , ...,  $1/s^n$ . That is,  $1/s$  is represented by a unit step function in time sequence form with intervals  $\Delta t$ ;  $\{\Delta t, \Delta t, \Delta t, \Delta t, \dots\}$ . We must remember that since this is representing  $\{h\}_A$ , the terms are the areas under the unit step curve. Similarly, for  $1/s^2$ , the time sequence consists of the areas under the unit ramp and so on.

In order to get a rational function for  $H(s)$  we use a slightly different form of polynomial to take into account the new operators. Let

$$H(s) = \frac{P(s)}{Q(s)} = \frac{a_0 + a_1/s + a_2/s^2 + \dots + a_n/s^n}{b_0 + b_1/s + b_2/s^2 + \dots + b_m/s^m}$$

$$= \frac{s^m}{s^n} \frac{a_0 s^n + a_1 s^{n-1} + a_2 s^{n-2} + \dots + a_n}{b_0 s^m + b_1 s^{m-1} + b_2 s^{m-2} + \dots + b_m} \quad (112)$$

For example, if the behavior of  $h(t)$  at  $t \rightarrow 0$  requires an  $H_1(s)$  whose denominator polynomial is 2 degrees higher than that of the numerator, we write

$$H_1(s) = \frac{a_2/s^2}{b_0 + b_1/s + b_2/s^2} \quad (113)$$

which is equivalent to

$$H_1(s) = \frac{a_2}{b_0 s^2 + b_1 s + b_2} \quad (114)$$

We then follow an entirely analogous procedure to that outlined in section 5.4. That is, we express Eq. 113 in time sequence form, using the time operator representations for  $1/s$ ,  $1/s^2$ , ...,  $1/s^n$ .

The capabilities of this method are equivalent to those of the previous method, and we may also use a hybrid system. In other words, we may write

$$H_1(s) = \frac{a_1/s}{cs + d + e/s} = \frac{a_1}{cs^2 + ds + e} \quad (115)$$

instead of Eq. 113 for the same situation, using the step, ramp, and higher integral time sequences for  $1/s$ ,  $1/s^2$ , ..., and the impulse, doublet, triplet, and higher order time sequences for  $1$ ,  $s$ ,  $s^2$ , and so on. Some test examples were worked out with these different representations. The results were identical. Since no new computational techniques were involved, the numerical work is not included here. The idea may be interesting because it opens up new approaches to the philosophy of time-domain operators and system functions.

## 5.6 CONCLUSION

We have shown that given a pair of input and output time functions, specified either in analytic form, graphically, or as time sequences of ordinate values at stated intervals, there exists a numerical procedure for obtaining the impulse response of the network that would give that response when excited by the given input. This procedure was obtained by converting the convolution integral into its component operations and using an iterative substitution method. This was replaced by a more compact process, namely, the synthetic division method for obtaining the time sequence of areas under the impulse response curve. Since it is the areas that really enter into the convolution process, it is sufficient to get a sequence of areas; and we indicated that these areas are correct to within the intervals of subdivision of the time abscissa. We also proved, by means of symbolical calculus and more completely by the Dirichlet series representation of the Laplace-Stieltjes integral for the Laplace transform, that this method is valid in producing the impulse response.

In discussing the question of the physical realizability of the network corresponding to that input-output pair, it was shown that two simple criteria can be easily applied to determine the answer. The first is evident by inspection; namely, to see if  $f_o(t)$  started before the excitation  $f_i(t)$  was applied. The other, which was used after  $h(t)$  passed the first test, is also easily discernible from the waveform of the impulse response obtained from that input-output pair.

Once the physical realizability of the network is established, we can go forward with the calculation of the system function  $H(s)$  from  $h(t)$ . The emphasis is on a simple straightforward procedure that utilizes only elementary algebra and is applicable to impulse responses that are given only graphically, or as time sequences of areas or ordinates. It can also be used for analytic functions, but we are not relying on that because it is not always possible to get  $h(t)$  in analytic form from  $f_i(t)$  and  $f_o(t)$ .

Three methods, all satisfying the requirements given above, were presented. Two of them depend on time domain operator representations for the system functions. One employs the doublets and triplets to represent the differentiator, double differentiator, and so on. The other uses the step functions, the ramp, and that group, to represent the operations of integration. These are of interest because of the picture they afford us of the mechanics of time operations. By their means, some illustrations can be given of phenomena for which the only explanation so far has been in the language of complex function theory; no doubt precise, but a little removed from the actual time domain in which the phenomenon is supposed to be taking place. These representations are accurate only if we keep  $\Delta t$  very small. This can well be appreciated if, for example, we look at a doublet. Because it is the area that remains constant, as we make  $\Delta t$  larger and larger,  $\{1/\Delta t, -1/\Delta t, 0, 0, 0, \dots\}$  looks less and less like a doublet.

On the other hand, the first method, described as the power series expansion method, is valid for all  $s$ , and because of the form of the series it assures good approximation

for  $t$  large. This we showed by means of the final value theorem. Also, because of the way we chose the relative degree of our numerator and denominator polynomials of  $H(s)$  we can guarantee the behavior of  $h(t)$  for  $t \rightarrow 0$ . Furthermore, if  $h(t)$  does have a rational function Laplace transform, we can get it precisely by matching a few of the beginning terms of the power series, sufficient to furnish the requisite number of linear, algebraic equations. All of the later terms are then automatically matched and the time function is recovered identically. If  $h(t)$  happens to have a transcendental Laplace transform, we can approximate the time function more and more accurately by using more and more of the terms of the power series in the determining equations. This follows from the fact that the successive coefficients of  $s$  in the power series are really shaping factors of the time curve. We illustrated these points by calculating  $H(s)$  from  $h(t)$  of both the rational function type and the transcendental type. We came out with good results, both in economy of computational effort and in economy of network elements.

## References

1. E. A. Guillemin, Communication Networks, Vol. II (John Wiley and Sons, New York, 1935).
2. O. Brune, J. Math. Phys. X, 191-236 (1931).
3. Y. W. Lee, J. Math. Phys. XI, 83-113 (1932).
4. C. E. M. Gewertz, Network Synthesis (Williams and Wilkins, Baltimore, 1933).
5. H. W. Bode, J. Math. Phys. XIII, 275-362 (1934).
6. E. L. Norton, Bell System Tech. J. 16, 178-193 (1937).
7. S. Darlington, J. Math. Phys. XVIII, 257-353 (1939).
8. W. Cauer, Elek. Nachr.-Tech. 16, 96-120 (1939).
9. E. A. Guillemin, A Summary of Modern Methods of Network Synthesis, Advances in Electronics, Vol. III (Academic Press, Inc., New York, 1951) pp. 261-303.
10. R. V. Churchill, Modern Operational Mathematics in Engineering (McGraw-Hill Book Company, Inc., New York, 1944).
11. W. B. Coulthard, Transients in Electric Circuits Using Heaviside Operational Calculus (Pitman, London, 1951).
12. M. F. Gardner, J. L. Barnes, Transients in Linear Systems (John Wiley and Sons, New York, 1942).
13. S. Goldman, Transformation Calculus and Electrical Transients (Prentice-Hall, Inc., New York, 1949).
14. N. W. McLachlan, Modern Operational Calculus (Macmillan and Company, Ltd., London, 1949).
15. B. van der Pol, H. Bremmer, Operational Calculus Based on the Two-Sided Laplace Integral (Cambridge University Press, England, 1950).
16. R. A. Fawwaz, A design procedure for pulse forming networks, M. Sc. Thesis, Department of Electrical Engineering, M.I.T., 1947.
17. R. B. Marsten, Synthesis of video amplifier for prescribed transient response, M. Sc. Thesis, Department of Electrical Engineering, M.I.T., 1946.
18. R. W. Bond, Synthesis of hydraulic acceleration servomechanisms, M. Sc. Thesis, Department of Electrical Engineering, M.I.T., 1947.
19. W. K. Linvill, Selection of the servomechanism transfer function from prescribed transient response, M. Sc. Thesis, M.I.T., 1945.
20. W. K. Linvill, Analysis of sampled data control systems, Doctoral Thesis, Department of Electrical Engineering, M.I.T., 1949.
21. E. A. Guillemin, A Historical Account of the Development of a Design Procedure for Pulse Forming Networks, Radiation Laboratory Report No. 43, M.I.T., Oct. 16, 1944.
22. G. N. Glasoe, J. V. Lebacqz, Pulse Generators, M.I.T. Radiation Laboratory Series, Vol. 5 (McGraw-Hill Book Company, Inc., New York, 1948).
23. L. Weinberg, New synthesis procedures for realizing transfer functions of RLC and RC networks, Technical Report No. 201, Research Laboratory of Electronics, M.I.T., Sept. 14, 1951.
24. D. F. Tuttle, Network synthesis for prescribed transient response, Doctoral Thesis, Department of Electrical Engineering, M.I.T., 1948.  
Baron Riche de Prony, Journal de l'École Polytechnique 2, 24-75 (Paris, 1794).
25. J. W. Carr, An analytic investigation of transient synthesis by exponentials, M. Sc. Thesis, Department of Electrical Engineering, M.I.T., 1949.

26. W. F. White, Application of transitional transients to network design, M. Sc. Thesis, Department of Electrical Engineering, M.I.T., 1950.
27. W. H. Kautz, A procedure for the synthesis of networks for specified transient response, Technical Report No. 209, Research Laboratory of Electronics, M.I.T., April 23, 1952.
28. M. V. Cerrillo, E. A. Guillemin, On basic existence theorems, Part III, Theoretical considerations on rational fraction expansions for network functions, Technical Report No. 233, Research Laboratory of Electronics, M.I.T., June 4, 1952.
29. H. M. Paynter, Methods and results from M.I.T. studies in unsteady flow, J. Boston Soc. Civil Engrs. 39, 120-165 (1952).
30. C. Jordan, Calculus of Finite Differences (Chelsea Publishing Co., New York, 1947).
31. L. M. Milne-Thomson, The Calculus of Finite Differences (Macmillan and Company, Ltd., London, 1951).
32. D. V. Widder, The Laplace Transform (Princeton University Press, 1946).
33. J. V. Uspensky, Introduction to Mathematical Probability (McGraw-Hill Book Company, Inc., New York, 1937).
34. E. C. Titchmarsh, The Theory of Functions (Oxford University Press, 1947).
35. W. W. Rogosinski, Volume and Integral (Interscience Publishers, New York, 1952).
36. T. J. Stieltjes, Jr., Annales de la Faculté des Sciences de Toulouse 8, 9 (1894, 1895).
37. C. J. de La Vallee Poussin, Leçons sur l'Approximation des Fonctions d'une Variable Réelle (Gauthier-Villars, Paris, 1919).
38. E. Jahnke, F. Emde, Tables of Functions with Formulae and Curves (Dover Publications, New York, 1945).
39. H. B. Dwight, Tables of Integrals and Other Mathematical Data (The Macmillan Company, New York, 1945).
40. H. E. Padé, Annales Scientifiques de l'École Normale Supérieure 9, 93 (1892); 24, 16 (1907).

